

SEQUENCE LISTING

<110> Coit, Doris  
Medina-Selby, Angelica  
Selby, Mark  
Houghton, Michael

<120> NOVEL HCV NON-STRUCTURAL POLYPEPTIDE

<130> PP01617.002

<140>

<141>

<160> 19

<170> PatentIn Ver. 2.0

<210> 1

<211> 9620

<212> DNA

<213> Artificial Sequence

<220>

<221> CDS

<222> (1990)..(7302)

<220>

<223> Description of Artificial Sequence: Hepatitis C pns345

<400> 1

cgcgcgttccgtatgacgttgaaaacctctgacacatgcagctccggagacggtcac 60

agcttgtctgttaaggatgcggggacagacaagccgtcagggcgcgtcagcggtgt 120

tggcgggtgtcggggctggcttaactatgcggcatcagagcagattgtac tgagagtgc 180

ccatatgaagcttttgc当地aaggcttaggcctccaaaaaagcctcctcac tacttctgga 240

atagctcaga ggccgaggcg gcctcggcctctgcataaataaaaaaatt agtcagccat 300

ggggcggaga atggcggaa ctggggggggaggaaattat tggctattgg ccattgcata 360

cgttgtatctatataatatgtacattt atattggctc atgtccaata tgaccgccat 420

gttgacatttattattgactatgtatataat agtaatcaat tacgggtca tttagttcata 480

gcccatatatggagttccgcgttacataac ttacggtaaa tggccgcctggctgaccgc 540

ccaacgacccccgcccatttacgtcaataat tgacgtatgttccatagta acgccaatag 600

ggactttcca ttgacgtcaa tgggtggagt atttacggta aactgcccac ttggcagtac 660

atcaagtgtatcatatgc当地a gtccgccccctattgacgtcaatgacggtaatggcccg 720

cctggcatta tgcccagttac atgaccttac gggactttcc tacttggcag tacatctacg 780

DRAFT 2/25/00

tattagtcat cgctattacc atggatgc gggtttggca gtacaccaat gggcgtggat 840  
agcggtttga ctcacgggaa tttccaagtc tccacccat tgacgtcaat gggagttgt 900  
tttggcacca aaatcaacgg gactttccaa aatgtcgtaa taacccgcc ccgttgacgc 960  
aaatggcg taggcgtgta cggtggagg tctataataag cagagctcgt ttagtgaacc 1020  
gtcagatcgc ctggagacgc catccacgct gtttgacct ccatagaaga caccgggacc 1080  
gatccagcct ccgcggccgg gaacggtgca ttggAACGCG gattccccgt gccaagagtg 1140  
acgtaagtac cgcctataga ctctataggc acacccctt ggcttttatg catgtatac 1200  
tgttttggc ttggggccta tacacccccc ctcttatgc tataggtgat ggtatagctt 1260  
agcctatagg tgtgggatat tgaccattat tgaccactcc cctattggtg acgataactt 1320  
ccattactaa tccataacat ggctttgc cacaactatc tctattggct atatgccaat 1380  
actctgtcct tcagagactg acacggactc tgtatTTTA caggatgggg tccattttt 1440  
atttacaaat tcacatatac aacaacgccc tccccgtgc ccgcagtttt tattaaacat 1500  
agcgtggat ctccgacatc tcgggtacgt gttccggaca tggcttttc tccggtagcg 1560  
gcggagcttc cacatccgag ccctggtccc atccgtccag cggctcatgg tcgctcggca 1620  
gctccttgct cctaacagtg gaggccagac ttaggcacag cacaatgccc accaccacca 1680  
gtgtccgca caaggccgtg gcggtagggt atgtgtctga aaatgagctc ggagattggg 1740  
ctcgcacctg gacgcagatg gaagacttaa ggcagcggca gaagaagatg caggcagctg 1800  
agttgttta ttctgataag agtcagaggt aactccgtt gcggtagctgt taacggtgaa 1860  
ggcagtgta gtctgagcag tactcggtc tgccgcgcg gccaccagac ataatagctg 1920  
acagactaac agactgttcc tttccatggg tctttctgc agtcaccgtc gtcgaccta 1980  
gaattcacc atg gct gca tat gca gct cag ggc tat aag gtg cta gta ctc 2031  
Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu  
1 5 10

aac ccc tct gtt gct gca aca ctg ggc ttt ggt gct tac atg tcc aag 2079  
Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys  
15 20 25 30

gct cat ggg atc gat cct aac atc agg acc ggg gtg aga aca att acc 2127  
Ala His Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr  
35 40 45

act ggc agc ccc atc acg tac tcc acc tac ggc aag ttc ctt gcc gac 2175  
Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp  
50 55 60

ggc ggg tgc tcg ggg ggc gct tat gac ata ata att tgt gac gag tgc 2223

Gly	Gly	Cys	Ser	Gly	Gly	Ala	Tyr	Asp	Ile	Ile	Ile	Cys	Asp	Glu	Cys	
65						70						75				
cac tcc acg gat gcc aca tcc atc ttg ggc att ggc act gtc ctt gac															2271	
His	Ser	Thr	Asp	Ala	Thr	Ser	Ile	Leu	Gly	Ile	Gly	Thr	Val	Leu	Asp	
80						85						90				
caa gca gag act gcg ggg gcg aga ctg gtt gtg ctc gcc acc gcc acc															2319	
Gln	Ala	Glu	Thr	Ala	Gly	Ala	Arg	Leu	Val	Val	Leu	Ala	Thr	Ala	Thr	
95						100						105			110	
cct ccg ggc tcc gtc act gtg ccc cat ccc aac atc gag gag gtt gct															2367	
Pro	Pro	Gly	Ser	Val	Thr	Val	Pro	His	Pro	Asn	Ile	Glu	Glu	Val	Ala	
115						120						125				
ctg tcc acc acc gga gag atc cct ttt tac ggc aag gct atc ccc ctc															2415	
Leu	Ser	Thr	Thr	Gly	Glu	Ile	Pro	Phe	Tyr	Gly	Lys	Ala	Ile	Pro	Leu	
130						135						140				
gaa gta atc aag ggg ggg aga cat ctc atc ttc tgt cat tca aag aag															2463	
Glu	Val	Ile	Lys	Gly	Gly	Arg	His	Leu	Ile	Phe	Cys	His	Ser	Lys	Lys	
145						150						155				
aag tgc gac gaa ctc gcc gca aag ctg gtc gca ttg ggc atc aat gcc															2511	
Lys	Cys	Asp	Glu	Leu	Ala	Ala	Lys	Leu	Val	Ala	Leu	Gly	Ile	Asn	Ala	
160						165						170				
gtg gcc tac tac cgc ggt ctt gac gtg tcc gtc atc ccg acc agc ggc															2559	
Val	Ala	Tyr	Tyr	Arg	Gly	Leu	Asp	Val	Ser	Val	Ile	Pro	Thr	Ser	Gly	
175						180						185			190	
gat gtt gtc gtc gtg gca acc gat ccc ctc atg acc ggc tat acc ggc															2607	
Asp	Val	Val	Val	Ala	Thr	Asp	Ala	Leu	Met	Thr	Gly	Tyr	Thr	Gly		
195						200						205				
gac ttc gac tcg gtg ata gac tgc aat acg tgt gtc acc cag aca gtc															2655	
Asp	Phe	Asp	Ser	Val	Ile	Asp	Cys	Asn	Thr	Cys	Val	Thr	Gln	Thr	Val	
210						215						220				
gat ttc agc ctt gac cct acc ttc acc att gag aca atc acg ctc ccc															2703	
Asp	Phe	Ser	Leu	Asp	Pro	Thr	Phe	Thr	Ile	Glu	Thr	Ile	Thr	Leu	Pro	
225						230						235				
caa gat gct tcc cgc act caa cgt cgg ggc agg act ggc agg ggg															2751	
Gln	Asp	Ala	Val	Ser	Arg	Thr	Gln	Arg	Arg	Gly	Arg	Thr	Gly	Arg	Gly	
240						245						250				
aag cca ggc atc tac aga ttt gtg gca ccg ggg gag cgc ccc tcc ggc															2799	
Lys	Pro	Gly	Ile	Tyr	Arg	Phe	Val	Ala	Pro	Gly	Glu	Arg	Pro	Ser	Gly	
255						260						265			270	
atg ttc gac tcg tcc gtc ctc tgt gag tgc tat gac gca ggc tgt gct															2847	
Met	Phe	Asp	Ser	Ser	Val	Leu	Cys	Glu	Cys	Tyr	Asp	Ala	Gly	Cys	Ala	
275						280						285				
tgg tat gag ctc acg ccc gcc gag act aca gtt agg cta cga gcg tac															2895	
Trp	Tyr	Glu	Leu	Thr	Pro	Ala	Glu	Thr	Val	Arg	Leu	Arg	Ala	Tyr		

DRAFT 2/2/90

290

295

300

atg aac acc ccg ggg ctt ccc gtg tgc cag gac cat ctt gaa ttt tgg Met Asn Thr Pro Gly Leu Pro Val Cys Gln Asp His Leu Glu Phe Trp 305	310	315	2943
gag ggc gtc ttt aca ggc ctc act cat ata gat gcc cac ttt cta tcc Glu Gly Val Phe Thr Gly Leu Thr His Ile Asp Ala His Phe Leu Ser 320	325	330	2991
cag aca aag cag agt ggg gag aac ctt cct tac ctg gta gcg tac caa Gln Thr Lys Gln Ser Gly Glu Asn Leu Pro Tyr Leu Val Ala Tyr Gln 335	340	345	3039
gcc acc gtg tgc gct agg gct caa gcc cct ccc cca tcg tgg gac cag Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser Trp Asp Gln 355	360	365	3087
atg tgg aag tgt ttg att cgc ctc aag ccc acc ctc cat ggg cca aca Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His Gly Pro Thr 370	375	380	3135
ccc ctg cta tac aga ctg ggc gct gtt cag aat gaa atc acc ctg acg Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Ile Thr Leu Thr 385	390	395	3183
cac cca gtc acc aaa tac atc atg aca tgc atg tcg gcc gac ctg gag His Pro Val Thr Lys Tyr Ile Met Thr Cys Met Ser Ala Asp Leu Glu 400	405	410	3231
gtc gtc acg agc acc tgg gtg ctc gtt ggc ggc gtc ctg gct gct ttg Val Val Thr Ser Thr Trp Val Leu Val Gly Gly Val Leu Ala Ala Leu 415	420	425	3279
gcc gcg tat tgc ctg tca aca ggc tgc gtg gtc ata gtg ggc agg gtc Ala Ala Tyr Cys Leu Ser Thr Gly Cys Val Val Ile Val Gly Arg Val 435	440	445	3327
gtc ttg tcc ggg aag ccg gca atc ata cct gac agg gaa gtc ctc tac Val Leu Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu Val Leu Tyr 450	455	460	3375
cga gag ttc gat gag atg gaa gag tgc tct cag cac tta ccg tac atc Arg Glu Phe Asp Glu Met Glu Glu Cys Ser Gln His Leu Pro Tyr Ile 465	470	475	3423
gag caa ggg atg atg ctc gcc gag cag ttc aag cag aag gcc ctc ggc Glu Gln Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys Ala Leu Gly 480	485	490	3471
ctc ctg cag acc gcg tcc cgt cag gca gag gtt atc gcc cct gct gtc Leu Leu Gln Thr Ala Ser Arg Gln Ala Glu Val Ile Ala Pro Ala Val 495	500	505	3519
cag acc aac tgg caa aaa ctc gag acc ttc tgg gcg aag cat atg tgg Gln Thr Asn Trp Gln Lys Leu Glu Thr Phe Trp Ala Lys His Met Trp 515	520	525	3567

0022222222222222

aac ttc atc agt ggg ata caa tac ttg gcg ggc ttg tca acg ctg cct		3615
Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro		
530	535	540
ggt aac ccc gcc att gct tca ttg atg gct ttt aca gct gct gtc acc		3663
Gly Asn Pro Ala Ile Ala Ser Leu Met Ala Phe Thr Ala Ala Val Thr		
545	550	555
agc cca cta acc act agc caa acc ctc ctc ttc aac ata ttg ggg ggg		3711
Ser Pro Leu Thr Thr Ser Gln Thr Leu Leu Phe Asn Ile Leu Gly Gly		
560	565	570
tgg gtg gct gcc cag ctc gcc ccc ggt gcc gct act gcc ttt gtg		3759
Trp Val Ala Ala Gln Leu Ala Ala Pro Gly Ala Ala Thr Ala Phe Val		
575	580	585
590		
ggc gct ggc tta gct ggc gcc atc ggc agt gtt gga ctg ggg aag		3807
Gly Ala Gly Leu Ala Gly Ala Ile Gly Ser Val Gly Leu Gly Lys		
595	600	605
gtc ctc ata gac atc ctt gca ggg tat ggc gcg ggc gtg gcg gga gct		3855
Val Leu Ile Asp Ile Leu Ala Gly Tyr Gly Ala Gly Val Ala Gly Ala		
610	615	620
ctt gtg gca ttc aag atc atg agc ggt gag gtc ccc tcc acg gag gac		3903
Leu Val Ala Phe Lys Ile Met Ser Gly Glu Val Pro Ser Thr Glu Asp		
625	630	635
ctg gtc aat cta ctg ccc gcc atc ctc tcg ccc gga gcc ctc gta gtc		3951
Leu Val Asn Leu Leu Pro Ala Ile Leu Ser Pro Gly Ala Leu Val Val		
640	645	650
ggc gtg gtc tgt gca gca ata ctg cgc cgg cac gtt ggc ccg ggc gag		3999
Gly Val Val Cys Ala Ala Ile Leu Arg Arg His Val Gly Pro Gly Glu		
655	660	665
670		
ggg gca gtg cag tgg atg aac cgg ctg ata gcc ttc gcc tcc cgg ggg		4047
Gly Ala Val Gln Trp Met Asn Arg Leu Ile Ala Phe Ala Ser Arg Gly		
675	680	685
aac cat gtt tcc ccc acg cac tac gtg ccg gag agc gat gca gct gcc		4095
Asn His Val Ser Pro Thr His Tyr Val Pro Glu Ser Asp Ala Ala Ala		
690	695	700
cgc gtc act gcc ata ctc agc agc ctc act gta acc cag ctc ctg agg		4143
Arg Val Thr Ala Ile Leu Ser Ser Leu Thr Val Thr Gln Leu Leu Arg		
705	710	715
cga ctg cac cag tgg ata agc tcg gag tgt acc act cca tgc tcc ggt		4191
Arg Leu His Gln Trp Ile Ser Ser Glu Cys Thr Thr Pro Cys Ser Gly		
720	725	730
tcc tgg cta agg gac atc tgg gac tgg ata tgc gag gtg ttg agc gac		4239
Ser Trp Leu Arg Asp Ile Trp Asp Trp Ile Cys Glu Val Leu Ser Asp		
735	740	745
750		
ttt aag acc tgg cta aaa gct aag ctc atg cca cag ctg cct ggg atc		4287

DRAFT 2/2/00

Phe	Lys	Thr	Trp	Leu	Lys	Ala	Lys	Leu	Met	Pro	Gln	Leu	Pro	Gly	Ile	
755								760							765	
ccc	ttt	gtg	tcc	tgc	cag	cgc	ggg	tat	aag	ggg	gtc	tgg	cga	ggg	gac	4335
Pro	Phe	Val	Ser	Cys	Gln	Arg	Gly	Tyr	Lys	Gly	Val	Trp	Arg	Gly	Asp	
770								775							780	
ggc	atc	atg	cac	act	cgc	tgc	cac	tgt	gga	gct	gag	atc	act	gga	cat	4383
Gly	Ile	Met	His	Thr	Arg	Cys	His	Cys	Gly	Ala	Glu	Ile	Thr	Gly	His	
785								790							795	
gtc	aaa	aac	ggg	acg	atg	agg	atc	gtc	ggt	cct	agg	acc	tgc	agg	aac	4431
Val	Lys	Asn	Gly	Thr	Met	Arg	Ile	Val	Gly	Pro	Arg	Thr	Cys	Arg	Asn	
800								805							810	
atg	tgg	agt	ggg	acc	ttc	ccc	att	aat	gcc	tac	acc	acg	ggc	ccc	tgt	4479
Met	Trp	Ser	Gly	Thr	Phe	Pro	Ile	Asn	Ala	Tyr	Thr	Gly	Pro	Cys		
815								820							825	830
acc	ccc	ctt	cct	gcg	ccg	aac	tac	acg	tcc	gcg	cta	tgg	agg	gtg	tct	4527
Thr	Pro	Leu	Pro	Ala	Pro	Asn	Tyr	Thr	Phe	Ala	Leu	Trp	Arg	Val	Ser	
								835			840				845	
gca	gag	gaa	tac	gtg	gag	ata	agg	cag	gtg	ggg	gac	ttc	cac	tac	gtg	4575
Ala	Glu	Glu	Tyr	Val	Glu	Ile	Arg	Gln	Val	Gly	Asp	Phe	His	Tyr	Val	
								850			855				860	
acg	ggt	atg	act	act	gac	aat	ctt	aaa	tgc	ccg	tgc	cag	gtc	cca	tcg	4623
Thr	Gly	Met	Thr	Thr	Asp	Asn	Leu	Lys	Cys	Pro	Cys	Gln	Val	Pro	Ser	
								865			870				875	
ccc	gaa	ttt	ttc	aca	gaa	ttg	gac	ggg	gtg	ccg	cta	cat	agg	ttt	gcg	4671
Pro	Glu	Phe	Phe	Thr	Glu	Leu	Asp	Gly	Val	Arg	Leu	His	Arg	Phe	Ala	
								880			885				890	
ccc	ccc	tgc	aag	ccc	ttg	ctg	cg	gg	gag	gta	tca	ttc	aga	gta	gga	4719
Pro	Pro	Cys	Lys	Pro	Leu	Leu	Arg	Glu	Glu	Val	Ser	Phe	Arg	Val	Gly	
								895			900				905	910
ctc	cac	gaa	tac	ccg	gta	ggg	tcg	caa	tta	cct	tgc	gag	ccc	gaa	ccg	4767
Leu	His	Glu	Tyr	Pro	Val	Gly	Ser	Gln	Leu	Pro	Cys	Glu	Pro	Glu	Pro	
								915			920				925	
gac	gtg	gcc	gtg	ttg	acg	tcc	atg	ctc	act	gat	ccc	tcc	cat	ata	aca	4815
Asp	Val	Ala	Val	Leu	Thr	Ser	Met	Leu	Thr	Asp	Pro	Ser	His	Ile	Thr	
								930			935				940	
gca	gag	gcf	gcc	ggg	cga	agg	ttg	gcf	agg	gga	tca	ccc	ccc	tct	gtg	4863
Ala	Glu	Ala	Ala	Gly	Arg	Arg	Leu	Ala	Arg	Gly	Ser	Pro	Pro	Ser	Val	
								945			950				955	
gcc	agc	tcc	tcg	gct	agc	cag	cta	tcc	gct	cca	tct	ctc	aag	gca	act	4911
Ala	Ser	Ser	Ser	Ala	Ser	Gln	Leu	Ser	Ala	Pro	Ser	Leu	Lys	Ala	Thr	
								960			965				970	
tgc	acc	gct	aac	cat	gac	tcc	cct	gat	gct	gag	ctc	ata	gag	gcc	aac	4959
Cys	Thr	Ala	Asn	His	Asp	Ser	Pro	Asp	Ala	Glu	Leu	Ile	Glu	Ala	Asn	

975                    980                    985                    990

ctc cta tgg agg cag gag atg ggc ggc aac atc acc agg gtt gag tca        5007  
 Leu Leu Trp Arg Gln Glu Met Gly Gly Asn Ile Thr Arg Val Glu Ser  
 995                    1000                    1005

gaa aac aaa gtg gtg att ctg gac tcc ttc gat ccg ctt gtg gcg gag        5055  
 Glu Asn Lys Val Val Ile Leu Asp Ser Phe Asp Pro Leu Val Ala Glu  
 1010                    1015                    1020

gag gac gag cg<sup>g</sup> gag atc tcc gta ccc gca gaa atc ctg cg<sup>g</sup> aag tct        5103  
 Glu Asp Glu Arg Glu Ile Ser Val Pro Ala Glu Ile Leu Arg Lys Ser  
 1025                    1030                    1035

cg<sup>g</sup> aga ttc gcc cag gcc ctg ccc gtt tgg gcg cg<sup>g</sup> gac tat aac        5151  
 Arg Arg Phe Ala Gln Ala Leu Pro Val Trp Ala Arg Pro Asp Tyr Asn  
 1040                    1045                    1050

ccc ccg cta gtg gag acg tgg aaa aag ccc gac tac gaa cca cct gtg        5199  
 Pro Pro Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr Glu Pro Pro Val  
 1055                    1060                    1065                    1070

gtc cat ggc tgc ccg ctt cca cct cca aag tcc cct cct gtg cct ccg        5247  
 Val His Gly Cys Pro Leu Pro Pro Lys Ser Pro Pro Val Pro Pro  
 1075                    1080                    1085

cct cg<sup>g</sup> aag aag cg<sup>g</sup> acg gtg gtc ctc act gaa tca acc cta tct act        5295  
 Pro Arg Lys Lys Arg Thr Val Val Leu Thr Glu Ser Thr Leu Ser Thr  
 1090                    1095                    1100

gcc ttg gcc gag ctc gcc acc aga agc ttt ggc agc tcc tca act tcc        5343  
 Ala Leu Ala Glu Leu Ala Thr Arg Ser Phe Gly Ser Ser Ser Thr Ser  
 1105                    1110                    1115

ggc att acg ggc gac aat acg aca aca tcc tct gag ccc gcc cct tct        5391  
 Gly Ile Thr Gly Asp Asn Thr Thr Ser Ser Glu Pro Ala Pro Ser  
 1120                    1125                    1130

ggc tgc ccc ccc gac tcc gac gct gag tcc tat tcc tcc atg ccc ccc        5439  
 Gly Cys Pro Pro Asp Ser Asp Ala Glu Ser Tyr Ser Ser Met Pro Pro  
 1135                    1140                    1145                    1150

ctg gag ggg gag cct ggg gat ccg gat ctt agc gac ggg tca tgg tca        5487  
 Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu Ser Asp Gly Ser Trp Ser  
 1155                    1160                    1165

acg gtc agt agt gag gcc aac gc<sup>g</sup> gag gat gtc gtg tgc tca atg        5535  
 Thr Val Ser Ser Glu Ala Asn Ala Glu Asp Val Val Cys Cys Ser Met  
 1170                    1175                    1180

tct tac tct tgg aca ggc gca ctc acc ccg tgc gcc gcg gaa gaa        5583  
 Ser Tyr Ser Trp Thr Gly Ala Leu Val Thr Pro Cys Ala Ala Glu Glu  
 1185                    1190                    1195

cag aaa ctg ccc atc aat gca cta agc aac tcg ttg cta cgt cac cac        5631  
 Gln Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu Arg His His  
 1200                    1205                    1210

DRAFT PEZ60

aat ttg gtg tat tcc acc acc tca cgc agt gct tgc caa agg cag aag			5679
Asn Leu Val Tyr Ser Thr Ser Arg Ser Ala Cys Gln Arg Gln Lys			
1215	1220	1225	1230
aaa gtc aca ttt gac aga ctg caa gtt ctg gac agc cat tac cag gac			5727
Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Ser His Tyr Gln Asp			
1235	1240	1245	
gta ctc aag gag gtt aaa gca gcg gcg tca aaa gtg aag gct aac ttg			5775
Val Leu Lys Glu Val Lys Ala Ala Ser Lys Val Lys Ala Asn Leu			
1250	1255	1260	
cta tcc gta gag gaa gct tgc agc ctg acg ccc cca cac tca gcc aaa			5823
Leu Ser Val Glu Glu Ala Cys Ser Leu Thr Pro Pro His Ser Ala Lys			
1265	1270	1275	
tcc aag ttt ggt tat ggg gca aaa gac gtc cgt tgc cat gcc aga aag			5871
Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Cys His Ala Arg Lys			
1280	1285	1290	
gcc gta acc cac atc aac tcc gtg tgg aaa gac ctt ctg gaa gac aat			5919
Ala Val Thr His Ile Asn Ser Val Trp Lys Asp Leu Leu Glu Asp Asn			
1295	1300	1305	1310
gta aca cca ata gac act acc atc atg gct aag aac gag gtt ttc tgc			5967
Val Thr Pro Ile Asp Thr Ile Met Ala Lys Asn Glu Val Phe Cys			
1315	1320	1325	
gtt cag cct gag aag ggg ggt cgt aag cca gct cgt ctc atc gtg ttc			6015
Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu Ile Val Phe			
1330	1335	1340	
ccc gat ctg ggc gtg cgc gtg tgc gaa aag atg gct ttg tac gac gtg			6063
Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu Tyr Asp Val			
1345	1350	1355	
gtt aca aag ctc ccc ttg gcc gtg atg gga agc tcc tac gga ttc caa			6111
Val Thr Lys Leu Pro Leu Ala Val Met Gly Ser Ser Tyr Gly Phe Gln			
1360	1365	1370	
tac tca cca gga cag cgg gtt gaa ttc ctc gtg caa gcg tgg aag tcc			6159
Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Gln Ala Trp Lys Ser			
1375	1380	1385	1390
aag aaa acc cca atg ggg ttc tcg tat gat acc cgc tgc ttt gac tcc			6207
Lys Lys Thr Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe Asp Ser			
1395	1400	1405	
aca gtc act gag agc gac atc cgt acg gag gag gca atc tac caa tgt			6255
Thr Val Thr Glu Ser Asp Ile Arg Thr Glu Glu Ala Ile Tyr Gln Cys			
1410	1415	1420	
tgt gac ctc gac ccc caa gcc cgc gtg gcc atc aag tcc ctc acc gag			6303
Cys Asp Leu Asp Pro Gln Ala Arg Val Ala Ile Lys Ser Leu Thr Glu			
1425	1430	1435	
agg ctt tat gtt ggg ggc cct ctt acc aat tca agg ggg gag aac tgc			6351

Arg	Leu	Tyr	Val	Gly	Gly	Pro	Leu	Thr	Asn	Ser	Arg	Gly	Glu	Asn	Cys	
1440				1445					1450							
ggc	tat	cgc	agg	tgc	cgc	gcf	agc	ggc	gta	ctg	aca	act	agc	tgt	ggt	6399
Gly	Tyr	Arg	Arg	Cys	Arg	Ala	Ser	Gly	Val	Leu	Thr	Thr	Ser	Cys	Gly	
1455				1460					1465				1470			
aac	acc	ctc	act	tgc	tac	atc	aag	gcc	cgg	qca	gcc	tgt	cga	gcc	gca	6447
Asn	Thr	Leu	Thr	Cys	Tyr	Ile	Lys	Ala	Arg	Ala	Ala	Cys	Arg	Ala	Ala	
1475						1480						1485				
ggg	ctc	cag	gac	tgc	acc	atg	ctc	gtg	tgt	ggc	gac	gac	tta	gtc	gtt	6495
Gly	Leu	Gln	Asp	Cys	Thr	Met	Leu	Val	Cys	Gly	Asp	Asp	Leu	Val	Val	
1490						1495							1500			
atc	tgt	gaa	agc	gcf	ggg	gtc	cag	gag	gac	gcf	gcf	agc	ctg	aga	gcc	6543
Ile	Cys	Glu	Ser	Ala	Gly	Val	Gln	Glu	Asp	Ala	Ala	Ser	Leu	Arg	Ala	
1505						1510						1515				
ttc	acg	gag	gct	atg	acc	agg	tac	tcc	gcc	ccc	cct	ggg	gac	ccc	cca	6591
Phe	Thr	Glu	Ala	Met	Thr	Arg	Tyr	Ser	Ala	Pro	Pro	Gly	Asp	Pro	Pro	
1520						1525						1530				
caa	cca	gaa	tac	gac	ttg	gag	ctc	ata	aca	tca	tgc	tcc	tcc	aac	gtg	6639
Gln	Pro	Glu	Tyr	Asp	Leu	Glu	Leu	Ile	Thr	Ser	Cys	Ser	Ser	Asn	Val	
1535						1540					1545			1550		
tca	gtc	gcc	cac	gac	ggc	gct	gga	aag	agg	gtc	tac	tac	ctc	acc	cgt	6687
Ser	Val	Ala	His	Asp	Gly	Ala	Gly	Lys	Arg	Val	Tyr	Tyr	Leu	Thr	Arg	
1555								1560					1565			
gac	cct	aca	acc	ccc	ctc	gcf	aga	gct	gcf	tgg	gag	aca	gca	aga	cac	6735
Asp	Pro	Thr	Thr	Pro	Leu	Ala	Arg	Ala	Ala	Trp	Glu	Thr	Ala	Arg	His	
1570								1575					1580			
act	cca	gtc	aat	tcc	tgg	cta	ggc	aac	ata	atc	atg	ttt	gcc	ccc	aca	6783
Thr	Pro	Val	Asn	Ser	Trp	Leu	Gly	Asn	Ile	Ile	Met	Phe	Ala	Pro	Thr	
1585						1590						1595				
ctg	tgg	gcf	agg	atg	ata	ctg	atg	acc	cat	tcc	ttt	agc	gtc	ctt	ata	6831
Leu	Trp	Ala	Arg	Met	Ile	Leu	Met	Thr	His	Phe	Phe	Ser	Val	Leu	Ile	
1600						1605						1610				
gcc	agg	gac	cag	ctt	gaa	cag	gdc	ctc	gat	tgc	gag	atc	tac	ggg	gcc	6879
Ala	Arg	Asp	Gln	Leu	Glu	Gln	Ala	Leu	Asp	Cys	Glu	Ile	Tyr	Gly	Ala	
1615						1620					1625			1630		
tgc	tac	tcc	ata	gaa	cca	ctg	gat	cta	cct	cca	atc	att	caa	aga	ctc	6927
Cys	Tyr	Ser	Ile	Glu	Pro	Leu	Asp	Leu	Pro	Pro	Ile	Ile	Gln	Arg	Leu	
1635								1640					1645			
cat	ggc	ctc	agc	gca	ttt	tca	ctc	cac	agt	tac	tct	cca	ggt	gaa	atc	6975
His	Gly	Leu	Ser	Ala	Phe	Ser	Leu	His	Ser	Tyr	Ser	Pro	Gly	Glu	Ile	
1650								1655					1660			
aat	agg	gtg	gcc	gca	tgc	ctc	aga	aaa	ctt	ggg	gta	ccg	ccc	ttg	cga	7023
Asn	Arg	Val	Ala	Ala	Cys	Leu	Arg	Lys	Leu	Gly	Val	Pro	Pro	Leu	Arg	

1665	1670	1675	
gct tgg aga cac cg <sup>g</sup> gcc cg <sup>g</sup> agc gtc cgc g <sup>c</sup> t agg ctt ctg gcc aga Ala Trp Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu Leu Ala Arg			7071
1680	1685	1690	
gga ggc agg gct gcc ata tgt ggc aag tac ctc ttc aac tgg gca gta Gly Gly Arg Ala Ala Ile Cys Gly Lys Tyr Leu Phe Asn Trp Ala Val			7119
1695	1700	1705	1710
aga aca aag ctc aaa ctc act cca ata gc <sup>g</sup> gc <sup>c</sup> g <sup>c</sup> t ggc cag ctg gac Arg Thr Lys Leu Lys Leu Thr Pro Ile Ala Ala Gly Gln Leu Asp			7167
1715	1720	1725	
ttg tcc ggc tgg ttc acg gct ggc tac agc g <sup>gg</sup> gga gac att tat cac Leu Ser Gly Trp Phe Thr Ala Gly Tyr Ser Gly Gly Asp Ile Tyr His			7215
1730	1735	1740	
agc gtg tct cat gcc cg <sup>g</sup> ccc cgc tgg atc tgg ttt tgc cta ctc ctg Ser Val Ser His Ala Arg Pro Arg Trp Ile Trp Phe Cys Leu Leu Leu			7263
1745	1750	1755	
ctt gct gca g <sup>gg</sup> gta ggc atc tac ctc ctc ccc aac cga tgaagg <sup>tt</sup> gg Leu Ala Ala Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg			7312
1760	1765	1770	
ggtaaacact ccggcctaaa aaaaaaaaaaa aatctagaaa gg <sup>cg</sup> cgccaa gatatacagg 7372			
atccactacg cg <sup>t</sup> tagagct cgctgatcag cctcgactgt gc <sup>tt</sup> tctagt tgccagccat 7432			
ctgttgtttg cccctccccc gtgccttcct tgaccctgga aggtgccact cccactgtcc 7492			
tttcctaata aaatgaggaa attgcatcgc attgtctgag taggtgtcat tctattctgg 7552			
ggggtgtgggt ggggcaggac agcaaggggg aggattggga agacaatagc aggcatgctg 7612			
gggagctctt ccgcttcctc gctcaactgac tc <sup>g</sup> ctgcgtc cggtcg <sup>tt</sup> cg gctgcggcga 7672			
gcggtatcag ctcactcaa ggcggtaata cg <sup>tt</sup> tatcca cagaatcagg ggataacgca 7732			
ggaaagaaca tgtgagcaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 7792			
ctggcg <sup>ttt</sup> tccataggct ccgccccct gacgagc <sup>at</sup> c aaaaaatcg acgctcaagt 7852			
cagaggtggc gaaacccgac aggactataa agataccagg cg <sup>ttt</sup> cccccc tggaaagctcc 7912			
ctcgtgcgtc ctcc <sup>t</sup> gttcc gaccctgccc cttaccggat acctgtccgc ctttctccct 7972			
tcgggaagcg tggcg <sup>ttt</sup> tcaatgctca cgctgttaggt atctcagttc ggtgttaggtc 8032			
gttcgctcca agctggctg tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta 8092			
tccggtaact atcgtcttga gtccaaacccg gtaagacacg acttatcgcc actggcagca 8152			
gccactggta acaggattag cagagcgagg tatgttaggcg gtgctacaga gttcttgaag 8212			
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 8272			

ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaac caccgctgg 8332  
agcggtggtt ttttggtt caagcagcag attacgcgcgca gaaaaaaagg atctcaagaa 8392  
gatccttga tctttctac ggggtctgac gctcagtggc acgaaaactc acgttaaggg 8452  
atttggtca tgagattatc aaaaaggatc ttcacctaga tcctttaaa taaaaaatga 8512  
agtttaat caatctaaag tatatatgag taaaacttggg ctgacagttt ccaatgctt 8572  
atcagtgagg cacctatctc agcgatctgt ctatccgtt catccatagt tgccctgactc 8632  
cccgtcggt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 8692  
ataccgcgag acccagcgtc accggctcca gatttatcga caataaacca gccagccgga 8752  
agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 8812  
tgccggaaag ctagagtaag tagttcgcca gttaatagtgt tgccaaacgt tggtgccatt 8872  
gctacaggca tcgtgggtgc acgctcgatcg tttggtatgg cttcattcag ctccgggttcc 8932  
caacgatcaa ggcgagttac atgatcccccc atggtgtgca aaaaagcgggt tagtccttc 8992  
ggtcctccga tcgttgtcag aagtaagttg gcccagttgt tatcactcat gtttatggca 9052  
gcactgcata attctttac tgtcatgcca tccgttaagat gctttctgt gactggtgag 9112  
tactcaacca agtcatctg agaatagtgt atgcggcgcac cgagttgctc ttgcccggcg 9172  
tcaatacggg ataataccgc gccacatagc agaactttaa aagtgtctcat cattggaaaa 9232  
cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttctgggtaa 9292  
cccactcgatc caccctactg atttcagca ttttttactt tcaccagcgt ttctgggtga 9352  
gcaaaaacag gaaggcaaaa tgccgaaaaa aagggaataaa gggcgacacg gaaatgttga 9412  
atactcatac tcttcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg 9472  
agcggataca tatttgaatg tatttagaaa aataaacaaa taggggttcc ggcacattt 9532  
ccccgaaaag tgccacctga cgtctaagaa accattatta tcatgacattt aacctataaa 9592  
aataggcgta tcacgaggcc ctggcgatc 9620

<210> 2  
<211> 1771  
<212> PRT  
<213> Hepatitis C virus

<220>

<223> Description of Artificial Sequence: Hepatitis C pns345

<400> 2

Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro

1                    5                    10                    15

Ser	Val	Ala	Ala	Thr	Leu	Gly	Phe	Gly	Ala	Tyr	Met	Ser	Lys	Ala	His
20															
25															
Gly	Ile	Asp	Pro	Asn	Ile	Arg	Thr	Gly	Val	Arg	Thr	Ile	Thr	Thr	Gly
35															
40															
45															
Ser	Pro	Ile	Thr	Tyr	Ser	Thr	Tyr	Gly	Lys	Phe	Leu	Ala	Asp	Gly	Gly
50															
55															
60															
Cys	Ser	Gly	Gly	Ala	Tyr	Asp	Ile	Ile	Ile	Cys	Asp	Glu	Cys	His	Ser
65															
70															
75															
80															
Thr	Asp	Ala	Thr	Ser	Ile	Leu	Gly	Ile	Gly	Thr	Val	Leu	Asp	Gln	Ala
85															
90															
95															
Glu	Thr	Ala	Gly	Ala	Arg	Leu	Val	Val	Leu	Ala	Thr	Ala	Thr	Pro	Pro
100															
105															
110															
Gly	Ser	Val	Thr	Val	Pro	His	Pro	Asn	Ile	Glu	Glu	Val	Ala	Leu	Ser
115															
120															
125															
Thr	Thr	Gly	Glu	Ile	Pro	Phe	Tyr	Gly	Lys	Ala	Ile	Pro	Leu	Glu	Val
130															
135															
140															
Ile	Lys	Gly	Gly	Arg	His	Leu	Ile	Phe	Cys	His	Ser	Lys	Lys	Lys	Cys
145															
150															
155															
160															
Asp	Glu	Leu	Ala	Ala	Lys	Leu	Val	Ala	Leu	Gly	Ile	Asn	Ala	Val	Ala
165															
170															
175															
Tyr	Tyr	Arg	Gly	Leu	Asp	Val	Ser	Val	Ile	Pro	Thr	Ser	Gly	Asp	Val
180															
185															
190															
Val	Val	Val	Ala	Thr	Asp	Ala	Leu	Met	Thr	Gly	Tyr	Thr	Gly	Asp	Phe
195															
200															
205															
Asp	Ser	Val	Ile	Asp	Cys	Asn	Thr	Cys	Val	Thr	Gln	Thr	Val	Asp	Phe
210															
215															
220															
Ser	Leu	Asp	Pro	Thr	Phe	Thr	Ile	Glu	Thr	Ile	Thr	Leu	Pro	Gln	Asp
225															
230															
235															
240															
Ala	Val	Ser	Arg	Thr	Gln	Arg	Gly	Arg	Thr	Gly	Arg	Gly	Lys	Pro	
245															
250															
255															
Gly	Ile	Tyr	Arg	Phe	Val	Ala	Pro	Gly	Glu	Arg	Pro	Ser	Gly	Met	Phe
260															
265															
270															
Asp	Ser	Ser	Val	Leu	Cys	Glu	Cys	Tyr	Asp	Ala	Gly	Cys	Ala	Trp	Tyr
275															
280															
285															
Glu	Leu	Thr	Pro	Ala	Glu	Thr	Thr	Val	Arg	Leu	Arg	Ala	Tyr	Met	Asn
290															
295															
300															
Thr	Pro	Gly	Leu	Pro	Val	Cys	Gln	Asp	His	Leu	Glu	Phe	Trp	Glu	Gly

305                    310                    315                    320  
 Val Phe Thr Gly Leu Thr His Ile Asp Ala His Phe Leu Ser Gln Thr  
 325                    330                    335  
 Lys Gln Ser Gly Glu Asn Leu Pro Tyr Leu Val Ala Tyr Gln Ala Thr  
 340                    345                    350  
 Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser Trp Asp Gln Met Trp  
 355                    360                    365  
 Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His Gly Pro Thr Pro Leu  
 370                    375                    380  
 Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Ile Thr Leu Thr His Pro  
 385                    390                    395                    400  
 Val Thr Lys Tyr Ile Met Thr Cys Met Ser Ala Asp Leu Glu Val Val  
 405                    410                    415  
 Thr Ser Thr Trp Val Leu Val Gly Gly Val Leu Ala Ala Leu Ala Ala  
 420                    425                    430  
 Tyr Cys Leu Ser Thr Gly Cys Val Val Ile Val Gly Arg Val Val Leu  
 435                    440                    445  
 Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu Val Leu Tyr Arg Glu  
 450                    455                    460  
 Phe Asp Glu Met Glu Glu Cys Ser Gln His Leu Pro Tyr Ile Glu Gln  
 465                    470                    475                    480  
 Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys Ala Leu Gly Leu Leu  
 485                    490                    495  
 Gln Thr Ala Ser Arg Gln Ala Glu Val Ile Ala Pro Ala Val Gln Thr  
 500                    505                    510  
 Asn Trp Gln Lys Leu Glu Thr Phe Trp Ala Lys His Met Trp Asn Phe  
 515                    520                    525  
 Ile Ser Gly Ile Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn  
 530                    535                    540  
 Pro Ala Ile Ala Ser Leu Met Ala Phe Thr Ala Ala Val Thr Ser Pro  
 545                    550                    555                    560  
 Leu Thr Thr Ser Gln Thr Leu Leu Phe Asn Ile Leu Gly Gly Trp Val  
 565                    570                    575  
 Ala Ala Gln Leu Ala Ala Pro Gly Ala Ala Thr Ala Phe Val Gly Ala  
 580                    585                    590  
 Gly Leu Ala Gly Ala Ala Ile Gly Ser Val Gly Leu Gly Lys Val Leu  
 595                    600                    605  
 Ile Asp Ile Leu Ala Gly Tyr Gly Ala Gly Val Ala Gly Ala Leu Val

610                    615                    620

Ala Phe Lys Ile Met Ser Gly Glu Val Pro Ser Thr Glu Asp Leu Val  
625                    630                    635                    640

Asn Leu Leu Pro Ala Ile Leu Ser Pro Gly Ala Leu Val Val Gly Val  
645                    650                    655

Val Cys Ala Ala Ile Leu Arg Arg His Val Gly Pro Gly Glu Gly Ala  
660                    665                    670

Val Gln Trp Met Asn Arg Leu Ile Ala Phe Ala Ser Arg Gly Asn His  
675                    680                    685

Val Ser Pro Thr His Tyr Val Pro Glu Ser Asp Ala Ala Ala Arg Val  
690                    695                    700

Thr Ala Ile Leu Ser Ser Leu Thr Val Thr Gln Leu Leu Arg Arg Leu  
705                    710                    715                    720

His Gln Trp Ile Ser Ser Glu Cys Thr Thr Pro Cys Ser Gly Ser Trp  
725                    730                    735

Leu Arg Asp Ile Trp Asp Trp Ile Cys Glu Val Leu Ser Asp Phe Lys  
740                    745                    750

Thr Trp Leu Lys Ala Lys Leu Met Pro Gln Leu Pro Gly Ile Pro Phe  
755                    760                    765

Val Ser Cys Gln Arg Gly Tyr Lys Gly Val Trp Arg Gly Asp Gly Ile  
770                    775                    780

Met His Thr Arg Cys His Cys Gly Ala Glu Ile Thr Gly His Val Lys  
785                    790                    795                    800

Asn Gly Thr Met Arg Ile Val Gly Pro Arg Thr Cys Arg Asn Met Trp  
805                    810                    815

Ser Gly Thr Phe Pro Ile Asn Ala Tyr Thr Thr Gly Pro Cys Thr Pro  
820                    825                    830

Leu Pro Ala Pro Asn Tyr Thr Phe Ala Leu Trp Arg Val Ser Ala Glu  
835                    840                    845

Glu Tyr Val Glu Ile Arg Gln Val Gly Asp Phe His Tyr Val Thr Gly  
850                    855                    860

Met Thr Thr Asp Asn Leu Lys Cys Pro Cys Gln Val Pro Ser Pro Glu  
865                    870                    875                    880

Phe Phe Thr Glu Leu Asp Gly Val Arg Leu His Arg Phe Ala Pro Pro  
885                    890                    895

Cys Lys Pro Leu Leu Arg Glu Glu Val Ser Phe Arg Val Gly Leu His  
900                    905                    910

Glu Tyr Pro Val Gly Ser Gln Leu Pro Cys Glu Pro Glu Pro Asp Val

915

920

925

Ala Val Leu Thr Ser Met Leu Thr Asp Pro Ser His Ile Thr Ala Glu  
 930 935 940

Ala Ala Gly Arg Arg Leu Ala Arg Gly Ser Pro Pro Ser Val Ala Ser  
 945 950 955 960

Ser Ser Ala Ser Gln Leu Ser Ala Pro Ser Leu Lys Ala Thr Cys Thr  
 965 970 975

Ala Asn His Asp Ser Pro Asp Ala Glu Leu Ile Glu Ala Asn Leu Leu  
 980 985 990

Trp Arg Gln Glu Met Gly Gly Asn Ile Thr Arg Val Glu Ser Glu Asn  
 995 1000 1005

Lys Val Val Ile Leu Asp Ser Phe Asp Pro Leu Val Ala Glu Glu Asp  
 1010 1015 1020

Glu Arg Glu Ile Ser Val Pro Ala Glu Ile Leu Arg Lys Ser Arg Arg  
 1025 1030 1035 1040

Phe Ala Gln Ala Leu Pro Val Trp Ala Arg Pro Asp Tyr Asn Pro Pro  
 1045 1050 1055

Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr Glu Pro Pro Val Val His  
 1060 1065 1070

Gly Cys Pro Leu Pro Pro Lys Ser Pro Pro Val Pro Pro Pro Arg  
 1075 1080 1085

Lys Lys Arg Thr Val Val Leu Thr Glu Ser Thr Leu Ser Thr Ala Leu  
 1090 1095 1100

Ala Glu Leu Ala Thr Arg Ser Phe Gly Ser Ser Ser Thr Ser Gly Ile  
 1095 1110 1115 1120

Thr Gly Asp Asn Thr Thr Ser Ser Glu Pro Ala Pro Ser Gly Cys  
 1125 1130 1135

Pro Pro Asp Ser Asp Ala Glu Ser Tyr Ser Ser Met Pro Pro Leu Glu  
 1140 1145 1150

Gly Glu Pro Gly Asp Pro Asp Leu Ser Asp Gly Ser Trp Ser Thr Val  
 1155 1160 1165

Ser Ser Glu Ala Asn Ala Glu Asp Val Val Cys Cys Ser Met Ser Tyr  
 1170 1175 1180

Ser Trp Thr Gly Ala Leu Val Thr Pro Cys Ala Ala Glu Glu Gln Lys  
 1185 1190 1195 1200

Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu Arg His His Asn Leu  
 1205 1210 1215

Val Tyr Ser Thr Thr Ser Arg Ser Ala Cys Gln Arg Gln Lys Lys Val

1220                    1225                    1230

Thr Phe Asp Arg Leu Gln Val Leu Asp Ser His Tyr Gln Asp Val Leu  
 1235                    1240                    1245

Lys Glu Val Lys Ala Ala Ala Ser Lys Val Lys Ala Asn Leu Leu Ser  
 1250                    1255                    1260

Val Glu Glu Ala Cys Ser Leu Thr Pro Pro His Ser Ala Lys Ser Lys  
 265                    1270                    1275                    1280

Phe Gly Tyr Gly Ala Lys Asp Val Arg Cys His Ala Arg Lys Ala Val  
 1285                    1290                    1295

Thr His Ile Asn Ser Val Trp Lys Asp Leu Leu Glu Asp Asn Val Thr  
 1300                    1305                    1310

Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu Val Phe Cys Val Gln  
 1315                    1320                    1325

Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu Ile Val Phe Pro Asp  
 1330                    1335                    1340

Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu Tyr Asp Val Val Thr  
 345                    1350                    1355                    1360

Lys Leu Pro Leu Ala Val Met Gly Ser Ser Tyr Gly Phe Gln Tyr Ser  
 1365                    1370                    1375

Pro Gly Gln Arg Val Glu Phe Leu Val Gln Ala Trp Lys Ser Lys Lys  
 1380                    1385                    1390

Thr Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe Asp Ser Thr Val  
 1395                    1400                    1405

Thr Glu Ser Asp Ile Arg Thr Glu Glu Ala Ile Tyr Gln Cys Cys Asp  
 1410                    1415                    1420

Leu Asp Pro Gln Ala Arg Val Ala Ile Lys Ser Leu Thr Glu Arg Leu  
 425                    1430                    1435                    1440

Tyr Val Gly Gly Pro Leu Thr Asn Ser Arg Gly Glu Asn Cys Gly Tyr  
 1445                    1450                    1455

Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr Ser Cys Gly Asn Thr  
 1460                    1465                    1470

Leu Thr Cys Tyr Ile Lys Ala Arg Ala Ala Cys Arg Ala Ala Gly Leu  
 1475                    1480                    1485

Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp Leu Val Val Ile Cys  
 1490                    1495                    1500

Glu Ser Ala Gly Val Gln Glu Asp Ala Ala Ser Leu Arg Ala Phe Thr  
 505                    1510                    1515                    1520

Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp Pro Pro Gln Pro

1525                    1530                    1535

Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser Asn Val Ser Val  
 1540                    1545                    1550

Ala His Asp Gly Ala Gly Lys Arg Val Tyr Tyr Leu Thr Arg Asp Pro  
 1555                    1560                    1565

Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr Ala Arg His Thr Pro  
 1570                    1575                    1580

Val Asn Ser Trp Leu Gly Asn Ile Ile Met Phe Ala Pro Thr Leu Trp  
 585                    1590                    1595                    1600

Ala Arg Met Ile Leu Met Thr His Phe Phe Ser Val Leu Ile Ala Arg  
 1605                    1610                    1615

Asp Gln Leu Glu Gln Ala Leu Asp Cys Glu Ile Tyr Gly Ala Cys Tyr  
 1620                    1625                    1630

Ser Ile Glu Pro Leu Asp Leu Pro Pro Ile Ile Gln Arg Leu His Gly  
 1635                    1640                    1645

Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro Gly Glu Ile Asn Arg  
 1650                    1655                    1660

Val Ala Ala Cys Leu Arg Lys Leu Gly Val Pro Pro Leu Arg Ala Trp  
 665                    1670                    1675                    1680

Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu Leu Ala Arg Gly Gly  
 1685                    1690                    1695

Arg Ala Ala Ile Cys Gly Lys Tyr Leu Phe Asn Trp Ala Val Arg Thr  
 1700                    1705                    1710

Lys Leu Lys Leu Thr Pro Ile Ala Ala Ala Gly Gln Leu Asp Leu Ser  
 1715                    1720                    1725

Gly Trp Phe Thr Ala Gly Tyr Ser Gly Gly Asp Ile Tyr His Ser Val  
 1730                    1735                    1740

Ser His Ala Arg Pro Arg Trp Ile Trp Phe Cys Leu Leu Leu Ala  
 745                    1750                    1755                    1760

Ala Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg  
 1765                    1770

<210> 3  
 <211> 9620  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> CDS  
 <222> (1990) .. (7302)

<220>

<223> Description of Artificial Sequence: pDeltaNS3NS5

<400> 3

cgcgcgtttc ggtgatgacg gtgaaaacct ctgacacatg cagctccgg agacggtcac 60  
agcttgtctg taagcgatg cgggagcag acaagccgt cagggcgcgt cagcgggtgt 120  
tggcgggtgt cggggctggc ttaactatgc ggcattcagag cagattgtac tgagagtgc 180  
ccatatgaag cttttgcaa aagcctaggc ctccaaaaaa gcctcctcac tacttctgga 240  
atagctcaga ggccgaggcg gcctcggcct ctgcataaat aaaaaaaaaatt agtcagccat 300  
ggggcggaga atggcggaa ctggcgggg agggattat tggctattgg ccattgcata 360  
cgttgtatct atatcataat atgtacattt atattggctc atgtccaata tgaccgccat 420  
gttgacattt attattgact agttattaaat agtaatcaat tacgggtca ttagttcata 480  
gccccatataat ggagttccgc gttacataac ttacggtaaa tggccgcct ggctgaccgc 540  
ccaacgaccc cgcgcattt acgtcaataa tgacgtatgt tcccatagta acgccaatag 600  
ggactttcca ttgacgtcaa tgggtggagt attacggta aactgcccac ttggcagtac 660  
atcaagtgtt tcatatgcca agtccgcccc ctattgacgt caatgacggt aaatggcccg 720  
cctggcatta tgcccagtac atgaccttac gggactttcc tacttggcag tacatctacg 780  
tattagtcat cgcttattacc atggtgatgc ggtttggca gtacaccaat gggcgtggat 840  
agcggtttga ctcacgggaa tttccaagtc tccacccat tgacgtcaat gggagttgt 900  
tttggcacca aaatcaacgg gactttccaa aatgtcgtaa taacccgcgc cgggtgacgc 960  
aaatggcgg taggcgtgtt cgggtggagg tctatataag cagagctcgt ttagtgaacc 1020  
gtcagatcgc ctggagacgc catccacgt gtttgaccc ccatagaaga caccggacc 1080  
gatccagcct ccgcggccgg gaacggtgca ttggAACGCG gattcccggt gccaaagagt 1140  
acgtaagtac cgcctataga ctctatagtc acacccctt ggctttagt catgtataac 1200  
tgtttttggc ttggggccta tacaccccg ctccatgtc tatagggtat ggtatagctt 1260  
agcctatagg tgtgggttat tgaccattat tgaccactcc cctattggtg acgatactt 1320  
ccattactaa tccataacat ggctttgc cacaactatc tctattggct atatgccaat 1380  
actctgtcct tcagagactg acacggactc tgtatTTTA caggatgggg tccattttt 1440  
atttacaaat tcacatatac aacaacgccc tccccgtgc ccgcagttt tattaaacat 1500  
agcgtggat ctccgacatc tcgggtacgt gttccggaca tggcttttc tccggtagcg 1560  
gcggagcttc cacatccgag ccctggtccc atccgtccag cggctcatgg tgcgtcggca 1620

gtccttgct cctaacagt gaggccagac ttaggcacag cacaatgccc accaccacca 1680  
 gtgtgccgca caaggccgtg gcggtagggt atgtgtctga aaatgagctc ggagattggg 1740  
 ctcgcacctg gacgcagatg gaagacttaa ggcagcggca gaagaagatg caggcagctg 1800  
 agttgttcta ttctgataag agtcagaggt aactcccgaa gcggtgctgt taacggtgaa 1860  
 gggcagtgtt gtcgtggcag tactcggtgc tgccgcgc gccaccagac ataatagctg 1920  
 acagactaac agactgttcc tttccatggg tctttctgc agtcaccgtc gtcgaccaa 1980  
 gaattcacc atg gct gca tat gca gct cag ggc tat aag gtg cta gta ctc 2031  
     Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu  
     1               5               10

aac	ccc	tct	gtt	gct	gca	aca	ctg	ggc	ttt	ggt	gct	tac	atg	tcc	aag	2079
Asn	Pro	Ser	Val	Ala	Ala	Thr	Leu	Gly	Phe	Gly	Ala	Tyr	Met	Ser	Lys	
15																30

gct	cat	ggg	atc	gat	cct	aac	atc	agg	acc	ggg	gtg	aga	aca	att	acc	2127
Ala	His	Gly	Ile	Asp	Pro	Asn	Ile	Arg	Thr	Gly	Val	Arg	Thr	Ile	Thr	
35																45

act	ggc	agc	ccc	atc	acg	tac	tcc	acc	tac	ggc	aag	ttc	ctt	gcc	gac	2175
Thr	Gly	Ser	Pro	Ile	Thr	Tyr	Ser	Thr	Tyr	Gly	Lys	Phe	Leu	Ala	Asp	
50																60

ggc	ggg	tgc	tcg	ggg	ggc	gct	tat	gac	ata	ata	att	tgt	gac	gag	tgc	2223
Gly	Gly	Cys	Ser	Gly	Gly	Ala	Tyr	Asp	Ile	Ile	Cys	Asp	Glu	Cys		
65																75

cac	tcc	acg	gat	gcc	aca	tcc	atc	ttg	ggc	att	ggc	act	gtc	ctt	gac	2271
His	Ser	Thr	Asp	Ala	Thr	Ser	Ile	Leu	Gly	Ile	Gly	Thr	Val	Leu	Asp	
80																90

caa	gca	gag	act	gcg	ggg	gcg	aga	ctg	gtt	gtg	ctc	gcc	acc	gcc	acc	2319
Gln	Ala	Glu	Thr	Ala	Gly	Ala	Arg	Leu	Val	Val	Leu	Ala	Thr	Ala	Thr	
95																110

cct	ccg	ggc	tcc	gtc	act	gtg	ccc	cat	ccc	aac	atc	gag	gag	gtt	gct	2367
Pro	Pro	Gly	Ser	Val	Thr	Val	Pro	His	Pro	Asn	Ile	Glu	Glu	Val	Ala	
115																125

ctg	tcc	acc	acc	gga	gag	atc	cct	ttt	tac	ggc	aag	gct	atc	ccc	ctc	2415
Leu	Ser	Thr	Thr	Gly	Glu	Ile	Pro	Phe	Tyr	Gly	Lys	Ala	Ile	Pro	Leu	
130																140

gaa	gta	atc	aag	ggg	ggg	aga	cat	ctc	atc	ttc	tgt	cat	tca	aag	aag	2463
Glu	Val	Ile	Lys	Gly	Gly	Arg	His	Leu	Ile	Phe	Cys	His	Ser	Lys	Lys	
145																155

aag	tgc	gac	gaa	ctc	gcc	gca	aag	ctg	gtc	gca	ttg	ggc	atc	aat	gcc	2511
Lys	Cys	Asp	Glu	Leu	Ala	Ala	Lys	Leu	Val	Ala	Leu	Gly	Ile	Asn	Ala	
160																170

gtg	gcc	tac	tac	cgc	ggt	ctt	gac	gtg	tcc	gtc	atc	ccg	acc	agc	ggc	2559
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------

Val	Ala	Tyr	Tyr	Arg	Gly	Leu	Asp	Val	Ser	Val	Ile	Pro	Thr	Ser	Gly	
175										185						190
gat gtt gtc gtc gtg gca acc gat gcc ctc atg acc ggc tat acc acc ggc															2607	
Asp	Val	Val	Val	Val	Ala	Thr	Asp	Ala	Leu	Met	Thr	Gly	Tyr	Thr	Gly	
										200						205
gac ttc gac tcg gtg ata gac tgc aat acg tgt gtc acc cag aca gtc															2655	
Asp	Phe	Asp	Ser	Val	Ile	Asp	Cys	Asn	Thr	Cys	Val	Thr	Gln	Thr	Val	
										215						220
gat ttc agc ctt gac cct acc ttc acc att gag aca atc acg ctc ccc															2703	
Asp	Phe	Ser	Leu	Asp	Pro	Thr	Phe	Thr	Ile	Glu	Thr	Ile	Thr	Leu	Pro	
										230						235
caa gat gct gtc tcc cgc actcaa cgt cgg ggc agg act ggc agg ggg															2751	
Gln	Asp	Ala	Val	Ser	Arg	Thr	Gln	Arg	Arg	Gly	Arg	Thr	Gly	Arg	Gly	
										245						250
aag cca ggc atc tac aga ttt gtg gca ccg ggg gag cgc ccc tcc ggc															2799	
Lys	Pro	Gly	Ile	Tyr	Arg	Phe	Val	Ala	Pro	Gly	Glu	Arg	Pro	Ser	Gly	
										255						260
atg ttc gac tcg tcc gtc ctc tgt gag tgc tat gac gca ggc tgt gct															2847	
Met	Phe	Asp	Ser	Ser	Val	Leu	Cys	Glu	Cys	Tyr	Asp	Ala	Gly	Cys	Ala	
										275						280
tgg tat gag ctc acg ccc gcc gag act aca gtt agg cta cga gcg tac															2895	
Trp	Tyr	Glu	Leu	Thr	Pro	Ala	Glu	Thr	Thr	Val	Arg	Leu	Arg	Ala	Tyr	
										290						295
atg aac acc ccg ggg ctt ccc gtg tgc cag gac cat ctt gaa ttt tgg															2943	
Met	Asn	Thr	Pro	Gly	Leu	Pro	Val	Cys	Gln	Asp	His	Leu	Glu	Phe	Trp	
										305						310
gag ggc gtc ttt aca ggc ctc act cat ata gat gcc cac ttt cta tcc															2991	
Glu	Gly	Val	Phe	Thr	Gly	Leu	Thr	His	Ile	Asp	Ala	His	Phe	Leu	Ser	
										320						325
cag aca aag cag agt ggg gag aac ctt cct tac ctg gta gcg tac caa															3039	
Gln	Thr	Lys	Gln	Ser	Gly	Glu	Asn	Leu	Pro	Tyr	Leu	Val	Ala	Tyr	Gln	
										335						340
gcc acc gtg tgc gct agg gct caa gcc cct ccc cca tcg tgg gac cag															3087	
Ala	Thr	Val	Cys	Ala	Arg	Ala	Gln	Ala	Pro	Pro	Pro	Ser	Trp	Asp	Gln	
										355						360
atg tgg aag tgt ttg att cgc ctc aag ccc acc ctc cat ggg cca aca															3135	
Met	Trp	Lys	Cys	Leu	Ile	Arg	Leu	Lys	Pro	Thr	Leu	His	Gly	Pro	Thr	
										370						375
ccc ctg cta tac aga ctg ggc gct gtt cag aat gaa atc acc ctg acg															3183	
Pro	Leu	Leu	Tyr	Arg	Leu	Gly	Ala	Val	Gln	Asn	Glu	Ile	Thr	Leu	Thr	
										385						390
cac cca gtc acc aaa tac atc atg aca tgc atg tcg gcc gac ctg gag															3231	
His	Pro	Val	Thr	Lys	Tyr	Ile	Met	Thr	Cys	Met	Ser	Ala	Asp	Leu	Glu	

400	405	410	
gtc gtc acg agc acc tgg gtg ctc gtt ggc ggc gtc ctg gct gct ttg Val Val Thr Ser Thr Trp Val Leu Val Gly Gly Val Leu Ala Ala Leu			3279
415	420	425	430
gcc gcg tat tgc ctg tca aca ggc tgc gtg gtc ata gtg ggc agg gtc Ala Ala Tyr Cys Leu Ser Thr Gly Cys Val Val Ile Val Gly Arg Val			3327
435	440		445
gtc ttg tcc ggg aag ccg gca atc ata cct gac agg gaa gtc ctc tac Val Leu Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu Val Leu Tyr			3375
450	455		460
cga gag ttc gat gag atg gaa gag tgc tct cag cac tta ccg tac atc Arg Glu Phe Asp Glu Met Glu Glu Cys Ser Gln His Leu Pro Tyr Ile			3423
465	470		475
gag caa ggg atg atg ctc gcc gag cag ttc aag cag aag gcc ctc ggc Glu Gln Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys Ala Leu Gly			3471
480	485		490
ctc ctg cag acc gcg tcc cgt cag gca gag gtt atc gcc cct gct gtc Leu Leu Gln Thr Ala Ser Arg Gln Ala Glu Val Ile Ala Pro Ala Val			3519
495	500		510
cag acc aac tgg caa aaa ctc gag acc ttc tgg gcg aag cat atg tgg Gln Thr Asn Trp Gln Lys Leu Glu Thr Phe Trp Ala Lys His Met Trp			3567
515	520		525
aac ttc atc agt ggg ata caa tac ttg gcg ggc ttg tca acg ctg cct Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro			3615
530	535		540
ggt aac ccc gcc att gct tca ttg atg gct ttt aca gct gct gtc acc Gly Asn Pro Ala Ile Ala Ser Leu Met Ala Phe Thr Ala Ala Val Thr			3663
545	550		555
agc cca cta acc act agc caa acc ctc ctc ttc aac ata ttg ggg ggg Ser Pro Leu Thr Thr Ser Gln Thr Leu Leu Phe Asn Ile Leu Gly Gly			3711
560	565		570
tgg gtg gct gcc cag ctc gcc ccc ggt gcc gct act gcc ttt gtg Trp Val Ala Ala Gln Leu Ala Ala Pro Gly Ala Ala Thr Ala Phe Val			3759
575	580		590
ggc gct ggc tta gct ggc gcc atc ggc agt gtt gga ctg ggg aag Gly Ala Gly Leu Ala Gly Ala Ile Gly Ser Val Gly Leu Gly Lys			3807
595	600		605
gtc ctc ata gac atc ctt gca ggg tat ggc gcg ggc gtg gcg gga gct Val Leu Ile Asp Ile Leu Ala Gly Tyr Gly Ala Gly Val Ala Gly Ala			3855
610	615		620
ctt gtg gca ttc aag atc atg agc ggt gag gtc ccc tcc acg gag gac Leu Val Ala Phe Lys Ile Met Ser Gly Glu Val Pro Ser Thr Glu Asp			3903
625	630		635

ctg gtc aat cta ctg ccc gcc atc ctc tcg ccc gga gcc ctc gta gtc		3951	
Leu Val Asn Leu Leu Pro Ala Ile Leu Ser Pro Gly Ala Leu Val Val			
640	645	650	
ggc gtg gtc tgt gca gca ata ctg cgc cg <sup>g</sup> cac gtt ggc ccg ggc gag		3999	
Gly Val Val Cys Ala Ala Ile Leu Arg Arg His Val Gly Pro Gly Glu			
655	660	665	670
ggg gca gtg cag tgg atg aac cg <sup>g</sup> ctg ata gcc ttc gcc tcc cg <sup>g</sup> ggg		4047	
Gly Ala Val Gln Trp Met Asn Arg Leu Ile Ala Phe Ala Ser Arg Gly			
675	680	685	
aac cat gtt tcc ccc acg cac tac gtg cg <sup>g</sup> gag agc gat gca gct gcc		4095	
Asn His Val Ser Pro Thr His Tyr Val Pro Glu Ser Asp Ala Ala Ala			
690	695	700	
cgc gtc act gcc ata ctc agc agc ctc act gta acc cag ctc ctg agg		4143	
Arg Val Thr Ala Ile Leu Ser Ser Leu Thr Val Thr Gln Leu Leu Arg			
705	710	715	
cga ctg cac cag tgg ata agc tcg gag tgt acc act cca tgc tcc ggt		4191	
Arg Leu His Gln Trp Ile Ser Ser Glu Cys Thr Thr Pro Cys Ser Gly			
720	725	730	
tcc tgg cta agg gac atc tgg gac tgg ata tgc gag gtg ttg agc gac		4239	
Ser Trp Leu Arg Asp Ile Trp Asp Trp Ile Cys Glu Val Leu Ser Asp			
735	740	745	750
ttt aag acc tgg cta aaa gct aag ctc atg cca cag ctg cct ggg atc		4287	
Phe Lys Thr Trp Leu Lys Ala Lys Leu Met Pro Gln Leu Pro Gly Ile			
755	760	765	
ccc ttt gtg tcc tgc cag cgc ggg tat aag ggg gtc tgg cga ggg gac		4335	
Pro Phe Val Ser Cys Gln Arg Gly Tyr Lys Gly Val Trp Arg Gly Asp			
770	775	780	
ggc atc atg cac act cgc tgc cac tgt gga gct gag atc act qga cat		4383	
Gly Ile Met His Thr Arg Cys His Cys Gly Ala Glu Ile Thr Gly His			
785	790	795	
gtc aaa aac ggg acg atg agg atc gtc ggt cct agg acc tgc agg aac		4431	
Val Lys Asn Gly Thr Met Arg Ile Val Gly Pro Arg Thr Cys Arg Asn			
800	805	810	
atg tgg agt ggg acc ttc ccc att aat gcc tac acc acg ggc ccc tgt		4479	
Met Trp Ser Gly Thr Phe Pro Ile Asn Ala Tyr Thr Thr Gly Pro Cys			
815	820	825	830
acc ccc ctt cct gcg ccg aac tac acg ttc gcg cta tgg agg gtg tct		4527	
Thr Pro Leu Pro Ala Pro Asn Tyr Thr Phe Ala Leu Trp Arg Val Ser			
835	840	845	
gca gag gaa tac gtg gag ata agg cag gtg ggg gac ttc cac tac gtg		4575	
Ala Glu Glu Tyr Val Glu Ile Arg Gln Val Gly Asp Phe His Tyr Val			
850	855	860	
acg ggt atg act act gac aat ctt aaa tgc ccg tgc cag gtc cca tcg		4623	

Thr Gly Met Thr Thr Asp Asn Leu Lys Cys Pro Cys Gln Val Pro Ser  
 865 870 875

ccc gaa ttt ttc aca gaa ttg gac ggg gtg cgc cta cat agg ttt gcg 4671  
 Pro Glu Phe Phe Thr Glu Leu Asp Gly Val Arg Leu His Arg Phe Ala  
 880 885 890

ccc ccc tgc aag ccc ttg ctg cgg gag gag gta tca ttc aga gta gga 4719  
 Pro Pro Cys Lys Pro Leu Leu Arg Glu Glu Val Ser Phe Arg Val Gly  
 895 900 905 910

ctc cac gaa tac ccg gta ggg tcg caa tta cct tgc gag ccc gaa ccg 4767  
 Leu His Glu Tyr Pro Val Gly Ser Gln Leu Pro Cys Glu Pro Glu Pro  
 915 920 925

gac gtg gcc gtg ttg acg tcc atg ctc act gat ccc tcc cat ata aca 4815  
 Asp Val Ala Val Leu Thr Ser Met Leu Thr Asp Pro Ser His Ile Thr  
 930 935 940

gca gag gcg gcc ggg cga agg ttg ggg agg gga tca ccc ccc tct gtg 4863  
 Ala Glu Ala Ala Gly Arg Arg Leu Ala Arg Gly Ser Pro Pro Ser Val  
 945 950 955

gcc agc tcc tcg gct agc cag cta tcc gct cca tct ctc aag gca act 4911  
 Ala Ser Ser Ser Ala Ser Gln Leu Ser Ala Pro Ser Leu Lys Ala Thr  
 960 965 970

tgc acc gct aac cat gac tcc cct gat gct gag ctc ata gag gcc aac 4959  
 Cys Thr Ala Asn His Asp Ser Pro Asp Ala Glu Leu Ile Glu Ala Asn  
 975 980 985 990

ctc cta tgg agg cag gag atg ggc ggc aac atc acc agg gtt gag tca 5007  
 Leu Leu Trp Arg Gln Glu Met Gly Gly Asn Ile Thr Arg Val Glu Ser  
 995 1000 1005

gaa aac aaa gtg gtg att ctg gac tcc ttc gat ccg ctt gtg gcg gag 5055  
 Glu Asn Lys Val Val Ile Leu Asp Ser Phe Asp Pro Leu Val Ala Glu  
 1010 1015 1020

gag gac gag cgg gag atc tcc gta ccc gca gaa atc ctg cgg aag tct 5103  
 Glu Asp Glu Arg Glu Ile Ser Val Pro Ala Glu Ile Leu Arg Lys Ser  
 1025 1030 1035

cg<sup>g</sup> aga ttc gcc cag gcc ctg ccc gtt tgg gcg cgg ccg gac tat aac 5151  
 Arg Arg Phe Ala Gln Ala Leu Pro Val Trp Ala Arg Pro Asp Tyr Asn  
 1040 1045 1050

ccc ccg cta gtg gag acg tgg aaa aag ccc gac tac gaa cca cct gtg 5199  
 Pro Pro Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr Glu Pro Pro Val  
 1055 1060 1065 1070

gtc cat ggc tgc ccg ctt cca cct cca aag tcc cct cct gtg cct ccg 5247  
 Val His Gly Cys Pro Leu Pro Pro Lys Ser Pro Pro Val Pro Pro  
 1075 1080 1085

cct ccg aag aag ccg acg gtg gtc ctc act gaa tca acc cta tct act 5295  
 Pro Arg Lys Lys Arg Thr Val Val Leu Thr Glu Ser Thr Leu Ser Thr

1090

1095

1100

gcc ttg gcc gag ctc gcc acc aga agc ttt ggc agc tcc tca act tcc Ala Leu Ala Glu Leu Ala Thr Arg Ser Phe Gly Ser Ser Ser Thr Ser 1105	1110	1115	5343	
ggc att acg ggc gac aat acg aca aca tcc tct gag ccc gcc cct tct Gly Ile Thr Gly Asp Asn Thr Thr Ser Ser Glu Pro Ala Pro Ser 1120	1125	1130	5391	
ggc tgc ccc ccc gac tcc gac gct gag tcc tat tcc tcc atg ccc ccc Gly Cys Pro Pro Asp Ser Asp Ala Glu Ser Tyr Ser Ser Met Pro Pro 1135	1140	1145	1150	5439
ctg gag ggg gag cct ggg gat ccg gat ctt agc gac ggg tca tgg tca Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu Ser Asp Gly Ser Trp Ser 1155	1160	1165	5487	
acg gtc agt agt gag gcc aac gcg gag gat gtc gtg tgc tgc tca atg Thr Val Ser Ser Glu Ala Asn Ala Glu Asp Val Val Cys Cys Ser Met 1170	1175	1180	5535	
tct tac tct tgg aca ggc gca ctc gtc acc ccg tgc gcc gcg gaa gaa Ser Tyr Ser Trp Thr Gly Ala Leu Val Thr Pro Cys Ala Ala Glu Glu 1185	1190	1195	5583	
cag aaa ctg ccc atc aat gca cta agc aac tcg ttg cta cgt cac cac Gln Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu Arg His His 1200	1205	1210	5631	
aat ttg gtg tat tcc acc acc tca cgc agt gct tgc caa agg cag aag Asn Leu Val Tyr Ser Thr Ser Arg Ser Ala Cys Gln Arg Gln Lys 1215	1220	1225	1230	5679
aaa gtc aca ttt gac aga ctg caa gtt ctg gac agc cat tac cag gac Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Ser His Tyr Gln Asp 1235	1240	1245	5727	
gta ctc aag gag gtt aaa gca gcg gcg tca aaa gtg aag gct aac ttg Val Leu Lys Glu Val Lys Ala Ala Ala Ser Lys Val Lys Ala Asn Leu 1250	1255	1260	5775	
cta tcc gta gag gaa gct tgc agt ctg acg ccc cca cac tca gcc aaa Leu Ser Val Glu Glu Ala Cys Ser Leu Thr Pro Pro His Ser Ala Lys 1265	1270	1275	5823	
tcc aag ttt ggt tat ggg gca aaa gac gtc cgt tgc cat gcc aga aag Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Cys His Ala Arg Lys 1280	1285	1290	5871	
gcc gta acc cac atc aac tcc gtg tgg aaa gac ctt ctg gaa gac aat Ala Val Thr His Ile Asn Ser Val Trp Lys Asp Leu Leu Glu Asp Asn 1295	1300	1305	1310	5919
gta aca cca ata gac act acc atc atg gct aag aac gag gtt ttc tgc Val Thr Pro Ile Asp Thr Ile Met Ala Lys Asn Glu Val Phe Cys 1315	1320	1325	5967	

gtt cag cct gag aag ggg ggt cgt aag cca gct cgt ctc atc gtg ttc		6015
Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu Ile Val Phe		
1330	1335	1340
ccc gat ctg ggc gtg cgc gtg tgc gaa aag atg gct ttg tac gac gtg		6063
Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu Tyr Asp Val		
1345	1350	1355
gtt aca aag ctc ccc ttg gcc gtg atg gga agc tcc tac gga ttc caa		6111
Val Thr Lys Leu Pro Leu Ala Val Met Gly Ser Ser Tyr Gly Phe Gln		
1360	1365	1370
tac tca cca gga cag cgg gtt gaa ttc ctc gtg caa gcg tgg aag tcc		6159
Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Gln Ala Trp Lys Ser		
1375	1380	1385
1390		
aag aaa acc cca atg ggg ttc tcg tat gat acc cgc tgc ttt gac tcc		6207
Lys Lys Thr Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe Asp Ser		
1395	1400	1405
aca gtc act gag agc gac atc cgt acg gag gag gca atc tac caa tgt		6255
Thr Val Thr Glu Ser Asp Ile Arg Thr Glu Glu Ala Ile Tyr Gln Cys		
1410	1415	1420
tgt gac ctc gac ccc caa gcc cgc gtg gdc atc aag tcc ctc acc gag		6303
Cys Asp Leu Asp Pro Gln Ala Arg Val Ala Ile Lys Ser Leu Thr Glu		
1425	1430	1435
agg ctt tat gtt ggg ggc cct ctt acc aat tca agg ggg gag aac tgc		6351
Arg Leu Tyr Val Gly Gly Pro Leu Thr Asn Ser Arg Gly Glu Asn Cys		
1440	1445	1450
ggc tat cgc agg tgc cgc gcg agc ggc gta ctg aca act agc tgt ggt		6399
Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr Ser Cys Gly		
1455	1460	1465
1470		
aac acc ctc act tgc tac atc aag gcc cgg gca gcc tgt cga gcc gca		6447
Asn Thr Leu Thr Cys Tyr Ile Lys Ala Arg Ala Ala Cys Arg Ala Ala		
1475	1480	1485
ggg ctc cag gac tgc acc atg ctc gtg tgt ggc gac gac tta gtc gtt		6495
Gly Leu Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp Leu Val Val		
1490	1495	1500
atc tgt gaa agc gcg ggg gtc cag gag gac gcg gcg agc ctg aga gcc		6543
Ile Cys Glu Ser Ala Gly Val Gln Glu Asp Ala Ala Ser Leu Arg Ala		
1505	1510	1515
ttc acg gag gct atg acc agg tac tcc gcc ccc cct ggg gac ccc cca		6591
Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp Pro Pro		
1520	1525	1530
caa cca gaa tac gac ttg gag ctc ata aca tca tgc tcc tcc aac gtg		6639
Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser Asn Val		
1535	1540	1545
1550		
tca gtc gcc cac gac ggc gct gga aag agg gtc tac tac ctc acc cgt		6687

Ser Val Ala His Asp Gly Ala Gly Lys Arg Val Tyr Tyr Leu Thr Arg  
 1555 1560 1565

gac cct aca acc ccc ctc gcg aga gct gcg tgg gag aca gca aga cac 6735  
 Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr Ala Arg His  
 1570 1575 1580

act cca gtc aat tcc tgg cta ggc aac ata atc atg ttt gcc ccc aca 6783  
 Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Phe Ala Pro Thr  
 1585 1590 1595

ctg tgg gcg agg atg ata ctg atg acc cat ttc ttt agc gtc ctt ata 6831  
 Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser Val Leu Ile  
 1600 1605 1610

gcc agg gac cag ctt gaa cag gcc ctc gat tgc gag atc tac ggg gcc 6879  
 Ala Arg Asp Gln Leu Glu Gln Ala Leu Asp Cys Glu Ile Tyr Gly Ala  
 1615 1620 1625 1630

tgc tac tcc ata gaa cca ctg gat cta cct cca atc att caa aga ctc 6927  
 Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro Pro Ile Ile Gln Arg Leu  
 1635 1640 1645

cat ggc ctc agc gca ttt tca ctc cac agt tac tct cca ggt gaa atc 6975  
 His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro Gly Glu Ile  
 1650 1655 1660

aat agg gtg gcc gca tgc ctc aga aaa ctt ggg gta ccg ccc ttg cga 7023  
 Asn Arg Val Ala Ala Cys Leu Arg Lys Leu Gly Val Pro Pro Leu Arg  
 1665 1670 1675

gct tgg aga cac cgg gcc cgg agc gtc cgc gct agg ctt ctg gcc aga 7071  
 Ala Trp Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu Leu Ala Arg  
 1680 1685 1690

gga ggc agg gct gcc ata tgt ggc aag tac ctc ttc aac tgg gca gta 7119  
 Gly Gly Arg Ala Ala Ile Cys Gly Lys Tyr Leu Phe Asn Trp Ala Val  
 1695 1700 1705 1710

aga aca aag ctc aaa ctc act cca ata gcg gcc gct ggc cag ctg gac 7167  
 Arg Thr Lys Leu Lys Leu Thr Pro Ile Ala Ala Gly Gln Leu Asp  
 1715 1720 1725

ttg tcc ggc tgg ttc acg gct ggc tac agc ggg gga gac att tat cac 7215  
 Leu Ser Gly Trp Phe Thr Ala Gly Tyr Ser Gly Gly Asp Ile Tyr His  
 1730 1735 1740

agc gtg tct cat gcc cgg ccc cgc tgg atc tgg ttt tgc cta ctc ctg 7263  
 Ser Val Ser His Ala Arg Pro Arg Trp Ile Trp Phe Cys Leu Leu Leu  
 1745 1750 1755

ctt gct gca ggg gta ggc atc tac ctc ccc aac cga tgaagggttgg 7312  
 Leu Ala Ala Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg  
 1760 1765 1770

ggtaaacact ccggcctaaa aaaaaaaaaaa aatcttagaaa ggccgcgccaa gatatcaagg 7372

atccactacg cgtagagct cgctgatcat cctcgactgt gccttctagt tgccagccat 7432  
ctgttgtttg cccctccccc gtgccttcct tgaccctgga aggtgccact cccactgtcc 7492  
ttcctaata aaatgaggaa attgcattgc attgtctgag taggtgtcat tctattctgg 7552  
gggtgggggt ggggcaggac agcaaggggg aggattggga agacaatagc aggcattgtg 7612  
ggagctctt ccgcttcctc gctcaactgac tcgctgcgtt cggtcgttcg gctgcggcga 7672  
gcgttatcat ctcactcaaa ggccgttaata cggttatcca cagaatcagg ggataacgca 7732  
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 7792  
ctggcggtttt tccataggct ccgcggccct gacgagdatc acaaaaatcg acgctcaagt 7852  
cagaggtggc gaaacccgac aggactataa agataccagg cggtttcccc tggaaagctcc 7912  
ctcgctcgct ctccgttcc gaccctgccc cttacggat acctgtccgc ctttctccct 7972  
tcgggaagcg tggcgcttcc tcaatgctca cgctgtaggt atctcagttc ggtgttaggtc 8032  
gttcgctcca agctggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta 8092  
tccggtaact atcgcttga gtccaaacccg gtaagacacg acttatcgcc actggcagca 8152  
gccactggta acaggattag cagagcgagg tatgttaggcg gtgctacaga gttttgaag 8212  
tggtggccta actacggcta cactagaagg acagtatttgc gatatctgcgc tctgctgaag 8272  
ccagttacct tcggaaaaag agttggtagc ttgtatccg gcaaacaac caccgctgg 8332  
agcggtggtt tttttgtttt caagcagcag attacgcgca gaaaaaaagg atctcaagaa 8392  
gatccttta tctttctac ggggtctgac gtcagtggaa acgaaaactc acgttaaggg 8452  
attttggta tgagattatc aaaaaggatc ttcacccatgaa tcctttaaa ttaaaaatga 8512  
agttttaat caatctaaag tatatatgag taaacttggt ctgacagttt ccaatgctta 8572  
atcagtgagg cacctatctc agcgatctgt ctatccgtt catccatagt tgcctgactc 8632  
cccgctgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 8692  
ataccgcgag acccacgctc accggctca gatttatcag caataaacca gccagccgga 8752  
agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 8812  
tgccggaaag ctagagtaag tagtcgcac gttaatagtt tgcgcaacgt tggtgcatt 8872  
gctacaggca tcgtgggtgc acgctcgatc tttggtatgg cttcatttcg ctccgggttcc 8932  
caacgatcaa ggcgagttac atgatcccc atgttgtgca aaaaagcggt tagtccttc 8992  
ggcctccga tcgttgcag aagtaagttg gcccgcgtgt tatcactcat gtttatggca 9052  
gcactgcata attcttac tgcgtatgcca tccgtaaat gctttctgt gactggtgag 9112

tactcaacca agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg 9172  
tcaatacggg ataataccgc gccacatagc agaacttaa aagtgctcat cattggaaaa 9232  
cgttcttcgg ggcgaaaact ctcaggatc ttaccgctgt tgagatccag ttcgatgtaa 9292  
cccactcgta cacccaactg atcttcagca tcctttactt tcaccagcgt ttctgggtga 9352  
gcaaaaacag gaaggcaaaa tgccgcaaaa aagggatcaa gggcacacg gaaatgttga 9412  
atactcatac tcttccttt tcaatattat tgaagcattt atcagggtta ttgtctcatg 9472  
agcggataca tatttgaatg tatttagaaa aataaadaaa taggggttcc ggcacattt 9532  
ccccgaaaag tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa 9592  
aataggcgta tcacgaggcc cttdcgta 9620

<210> 4  
<211> 1771  
<212> PRT  
<213> Artificial Sequence

<400> 4  
Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro  
1 5 10 15  
Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His  
20 25 30  
Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly  
35 40 45  
Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly  
50 55 60  
Cys Ser Gly Gly Ala Tyr Asp Ile Ile Cys Asp Glu Cys His Ser  
65 70 75 80  
Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Ala  
85 90 95  
Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala Thr Ala Thr Pro Pro  
100 105 110  
Gly Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser  
115 120 125  
Thr Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val  
130 135 140  
Ile Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Lys Lys Cys  
145 150 155 160  
Asp Glu Leu Ala Ala Lys Leu Val Ala Leu Gly Ile Asn Ala Val Ala  
165 170 175

Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val  
 180 185 190  
 Val Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe  
 195 200 205  
 Asp Ser Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe  
 210 215 220  
 Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr Ile Thr Leu Pro Gln Asp  
 225 230 235 240  
 Ala Val Ser Arg Thr Gln Arg Arg Gly Arg Thr Gly Arg Gly Lys Pro  
 245 250 255  
 Gly Ile Tyr Arg Phe Val Ala Pro Gly Glu Arg Pro Ser Gly Met Phe  
 260 265 270  
 Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp Ala Gly Cys Ala Trp Tyr  
 275 280 285  
 Glu Leu Thr Pro Ala Glu Thr Thr Val Arg Leu Arg Ala Tyr Met Asn  
 290 295 300  
 Thr Pro Gly Leu Pro Val Cys Gln Asp His Leu Glu Phe Trp Glu Gly  
 305 310 315 320  
 Val Phe Thr Gly Leu Thr His Ile Asp Ala His Phe Leu Ser Gln Thr  
 325 330 335  
 Lys Gln Ser Gly Glu Asn Leu Pro Tyr Leu Val Ala Tyr Gln Ala Thr  
 340 345 350  
 Val Cys Ala Arg Ala Gln Ala Pro Pro Ser Trp Asp Gln Met Trp  
 355 360 365  
 Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His Gly Pro Thr Pro Leu  
 370 375 380  
 Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Ile Thr Leu Thr His Pro  
 385 390 395 400  
 Val Thr Lys Tyr Ile Met Thr Cys Met Ser Ala Asp Leu Glu Val Val  
 405 410 415  
 Thr Ser Thr Trp Val Leu Val Gly Gly Val Leu Ala Ala Leu Ala Ala  
 420 425 430  
 Tyr Cys Leu Ser Thr Gly Cys Val Val Ile Val Gly Arg Val Val Leu  
 435 440 445  
 Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu Val Leu Tyr Arg Glu  
 450 455 460  
 Phe Asp Glu Met Glu Glu Cys Ser Gln His Leu Pro Tyr Ile Glu Gln  
 465 470 475 480

Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys Ala Leu Gly Leu Leu  
 485 490 495

Gln Thr Ala Ser Arg Gln Ala Glu Val Ile Ala Pro Ala Val Gln Thr  
 500 505 510

Asn Trp Gln Lys Leu Glu Thr Phe Trp Ala Lys His Met Trp Asn Phe  
 515 520 525

Ile Ser Gly Ile Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn  
 530 535 540

Pro Ala Ile Ala Ser Leu Met Ala Phe Thr Ala Ala Val Thr Ser Pro  
 545 550 555 560

Leu Thr Thr Ser Gln Thr Leu Leu Phe Asn Ile Leu Gly Gly Trp Val  
 565 570 575

Ala Ala Gln Leu Ala Ala Pro Gly Ala Ala Thr Ala Phe Val Gly Ala  
 580 585 590

Gly Leu Ala Gly Ala Ala Ile Gly Ser Val Gly Leu Gly Lys Val Leu  
 595 600 605

Ile Asp Ile Leu Ala Gly Tyr Gly Ala Gly Val Ala Gly Ala Leu Val  
 610 615 620

Ala Phe Lys Ile Met Ser Gly Glu Val Pro Ser Thr Glu Asp Leu Val  
 625 630 635 640

Asn Leu Leu Pro Ala Ile Leu Ser Pro Gly Ala Leu Val Val Gly Val  
 645 650 655

Val Cys Ala Ala Ile Leu Arg Arg His Val Gly Pro Gly Glu Gly Ala  
 660 665 670

Val Gln Trp Met Asn Arg Leu Ile Ala Phe Ala Ser Arg Gly Asn His  
 675 680 685

Val Ser Pro Thr His Tyr Val Pro Glu Ser Asp Ala Ala Ala Arg Val  
 690 695 700

Thr Ala Ile Leu Ser Ser Leu Thr Val Thr Gln Leu Leu Arg Arg Leu  
 705 710 715 720

His Gln Trp Ile Ser Ser Glu Cys Thr Thr Pro Cys Ser Gly Ser Trp  
 725 730 735

Leu Arg Asp Ile Trp Asp Trp Ile Cys Glu Val Leu Ser Asp Phe Lys  
 740 745 750

Thr Trp Leu Lys Ala Lys Leu Met Pro Gln Leu Pro Gly Ile Pro Phe  
 755 760 765

Val Ser Cys Gln Arg Gly Tyr Lys Gly Val Trp Arg Gly Asp Gly Ile  
 770 775 780

Met His Thr Arg Cys His Cys Gly Ala Glu Ile Thr Gly His Val Lys  
785 790 795 800

Asn Gly Thr Met Arg Ile Val Gly Pro Arg Thr Cys Arg Asn Met Trp  
805 810 815

Ser Gly Thr Phe Pro Ile Asn Ala Tyr Thr Thr Gly Pro Cys Thr Pro  
820 825 830

Leu Pro Ala Pro Asn Tyr Thr Phe Ala Leu Trp Arg Val Ser Ala Glu  
835 840 845

Glu Tyr Val Glu Ile Arg Gln Val Gly Asp Phe His Tyr Val Thr Gly  
850 855 860

Met Thr Thr Asp Asn Leu Lys Cys Pro Cys Gln Val Pro Ser Pro Glu  
865 870 875 880

Phe Phe Thr Glu Leu Asp Gly Val Arg Leu His Arg Phe Ala Pro Pro  
885 890 895

Cys Lys Pro Leu Leu Arg Glu Glu Val Ser Phe Arg Val Gly Leu His  
900 905 910

Glu Tyr Pro Val Gly Ser Gln Leu Pro Cys Glu Pro Glu Pro Asp Val  
915 920 925

Ala Val Leu Thr Ser Met Leu Thr Asp Pro Ser His Ile Thr Ala Glu  
930 935 940

Ala Ala Gly Arg Arg Leu Ala Arg Gly Ser Pro Pro Ser Val Ala Ser  
945 950 955 960

Ser Ser Ala Ser Gln Leu Ser Ala Pro Ser Leu Lys Ala Thr Cys Thr  
965 970 975

Ala Asn His Asp Ser Pro Asp Ala Glu Leu Ile Glu Ala Asn Leu Leu  
980 985 990

Trp Arg Gln Glu Met Gly Gly Asn Ile Thr Arg Val Glu Ser Glu Asn  
995 1000 1005

Lys Val Val Ile Leu Asp Ser Phe Asp Pro Leu Val Ala Glu Glu Asp  
1010 1015 1020

Glu Arg Glu Ile Ser Val Pro Ala Glu Ile Leu Arg Lys Ser Arg Arg  
1025 1030 1035 1040

Phe Ala Gln Ala Leu Pro Val Trp Ala Arg Pro Asp Tyr Asn Pro Pro  
1045 1050 1055

Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr Glu Pro Pro Val Val His  
1060 1065 1070

Gly Cys Pro Leu Pro Pro Pro Lys Ser Pro Pro Val Pro Pro Pro Arg  
1075 1080 1085

Lys Lys Arg Thr Val Val Leu Thr Glu Ser Thr Leu Ser Thr Ala Leu  
1090 1095 1100

Ala Glu Leu Ala Thr Arg Ser Phe Gly Ser Ser Ser Thr Ser Gly Ile  
105 1110 1115 1120

Thr Gly Asp Asn Thr Thr Ser Ser Glu Pro Ala Pro Ser Gly Cys  
1125 1130 1135

Pro Pro Asp Ser Asp Ala Glu Ser Tyr Ser Ser Met Pro Pro Leu Glu  
1140 1145 1150

Gly Glu Pro Gly Asp Pro Asp Leu Ser Asp Gly Ser Trp Ser Thr Val  
1155 1160 1165

Ser Ser Glu Ala Asn Ala Glu Asp Val Val Cys Cys Ser Met Ser Tyr  
1170 1175 1180

Ser Trp Thr Gly Ala Leu Val Thr Pro Cys Ala Ala Glu Glu Gln Lys  
185 1190 1195 1200

Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu Arg His His Asn Leu  
1205 1210 1215

Val Tyr Ser Thr Thr Ser Arg Ser Ala Cys Gln Arg Gln Lys Lys Val  
1220 1225 1230

Thr Phe Asp Arg Leu Gln Val Leu Asp Ser His Tyr Gln Asp Val Leu  
1235 1240 1245

Lys Glu Val Lys Ala Ala Ser Lys Val Lys Ala Asn Leu Leu Ser  
1250 1255 1260

Val Glu Glu Ala Cys Ser Leu Thr Pro Pro His Ser Ala Lys Ser Lys  
265 1270 1275 1280

Phe Gly Tyr Gly Ala Lys Asp Val Arg Cys His Ala Arg Lys Ala Val  
1285 1290 1295

Thr His Ile Asn Ser Val Trp Lys Asp Leu Leu Glu Asp Asn Val Thr  
1300 1305 1310

Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu Val Phe Cys Val Gln  
1315 1320 1325

Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu Ile Val Phe Pro Asp  
1330 1335 1340

Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu Tyr Asp Val Val Thr  
345 1350 1355 1360

Lys Leu Pro Leu Ala Val Met Gly Ser Ser Tyr Gly Phe Gln Tyr Ser  
1365 1370 1375

Pro Gly Gln Arg Val Glu Phe Leu Val Gln Ala Trp Lys Ser Lys Lys  
1380 1385 1390

Thr Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe Asp Ser Thr Val  
1395 1400 1405

Thr Glu Ser Asp Ile Arg Thr Glu Glu Ala Ile Tyr Gln Cys Cys Asp  
1410 1415 1420

Leu Asp Pro Gln Ala Arg Val Ala Ile Lys Ser Leu Thr Glu Arg Leu  
425 1430 1435 1440

Tyr Val Gly Gly Pro Leu Thr Asn Ser Arg Gly Glu Asn Cys Gly Tyr  
1445 1450 1455

Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr Ser Cys Gly Asn Thr  
1460 1465 1470

Leu Thr Cys Tyr Ile Lys Ala Arg Ala Ala Cys Arg Ala Ala Gly Leu  
1475 1480 1485

Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp Leu Val Val Ile Cys  
1490 1495 1500

Glu Ser Ala Gly Val Gln Glu Asp Ala Ala Ser Leu Arg Ala Phe Thr  
505 1510 1515 1520

Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp Pro Pro Gln Pro  
1525 1530 1535

Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser Asn Val Ser Val  
1540 1545 1550

Ala His Asp Gly Ala Gly Lys Arg Val Tyr Tyr Leu Thr Arg Asp Pro  
1555 1560 1565

Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr Ala Arg His Thr Pro  
1570 1575 1580

Val Asn Ser Trp Leu Gly Asn Ile Ile Met Phe Ala Pro Thr Leu Trp  
585 1590 1595 1600

Ala Arg Met Ile Leu Met Thr His Phe Phe Ser Val Leu Ile Ala Arg  
1605 1610 1615

Asp Gln Leu Glu Gln Ala Leu Asp Cys Glu Ile Tyr Gly Ala Cys Tyr  
1620 1625 1630

Ser Ile Glu Pro Leu Asp Leu Pro Pro Ile Ile Gln Arg Leu His Gly  
1635 1640 1645

Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro Gly Glu Ile Asn Arg  
1650 1655 1660

Val Ala Ala Cys Leu Arg Lys Leu Gly Val Pro Pro Leu Arg Ala Trp  
665 1670 1675 1680

Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu Leu Ala Arg Gly Gly  
1685 1690 1695

Arg Ala Ala Ile Cys Gly Lys Tyr Leu Phe Asn Trp Ala Val Arg Thr  
1700 1705 1710

Lys Leu Lys Leu Thr Pro Ile Ala Ala Ala Gly Gln Leu Asp Leu Ser  
1715 1720 1725

Gly Trp Phe Thr Ala Gly Tyr Ser Gly Gly Asp Ile Tyr His Ser Val  
1730 1735 1740

Ser His Ala Arg Pro Arg Trp Ile Trp Phe Cys Leu Leu Leu Leu Ala  
745 1750 1755 1760

Ala Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg  
1765 1770

<210> 5

<211> 4282

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: pCMVII

<400> 5

tcgcgcgttt cggatgac ggtaaaaacc tctgacacat gcagctcccg gagacggta 60  
cagcttgcgt gtaaggcgat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120  
ttggcgggtg tcggggctgg cttaactatg cggcattcaga gcagattgta ctgagagtgc 180  
accatatgaa gcttttgca aaagcctagg cctcaaaaaa agcctcctca ctacttctgg 240  
aatagctcag aggccgaggc gcctcggcc tctgcataaa taaaaaaaaat tagtcagcca 300  
tggggcggag aatgggcgga actgggcggg gagggattt ttggctattt gcccatttcat 360  
acgttgtatc tatatcataa tatgtacatt tatattggct catgtccaaat atgaccgcca 420  
tgttgcattt gattattgac tagttattaa tagtaatcaa ttacgggttc attagttcat 480  
agcccatata tggagttccg cgttacataa cttacggtaa atggccgcc tggctgaccg 540  
cccaacgacc cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata 600  
gggactttcc attgacgtca atgggtggag tattacggt aaactgcccc cttggcagta 660  
catcaagtgt atcatatgcc aagtccgccc cttattgacg tcaatgacgg taaatggccc 720  
gcctggcatt atgcccagta catgacccta cgggactttc ctacttggca gtacatctac 780  
gtattagtca tcgctattac catggtgatg cgggtttggc agtacacccaa tggcgtgga 840  
tagcggtttgc actcacgggg atttccaaatg ctccacccca ttgacgtcaa tggagtttgc 900  
ttttggcacc aaaatcaacg ggactttcca aaatgtcgta ataacccgc cccgttgacg 960

DNA sequence

caaatggcg gttaggcgtgt acgggtggag gtcttatataa gdagagctcg ttttagtgaac 1020  
cgtcagatcg cctggagacg ccatccacgc tgtttgacc tccatagaag acaccggac 1080  
cgatccagcc tccgcggccg ggaacgggtgc attggaacgc ggattccccg tgccaagagt 1140  
gacgtaagta ccgcctatag actctatagg cacacccctt tggctcttat gcatgctata 1200  
ctgttttgg cttggggcct atacacccccc gtttccttat gctataggtg atggtatacg 1260  
ttagcctata ggtgtgggtt attgaccatt attgaccact cccctattgg tgacgatact 1320  
ttccattact aatccataac atggctctt gccacaacta tctctattgg ctatatgcc 1380  
atactctgtc cttcagagac tgacacggac tctgtatttt tacaggatgg ggtcccattt 1440  
attatttaca aattcacata tacaacaacg ccgtcccccg tgcccgcaagt ttttattaaa 1500  
catagcgtgg gatctccacg cgaatctcggtacgtgttc cggacatggg ctcttcctcg 1560  
gtagcggcgg agcttccaca tccgagccct ggtcccatgc ctccagcggc tcattggcgc 1620  
tcggcagctc cttgctccta acagtggagg ccagacttag gcacagcaca atgcccacca 1680  
ccaccagtgt gccgcacaag gccgtggcgg taggttatgt gtctgaaaat gagctcggag 1740  
attgggctcg caccgctgac gcagatggaa gactaaggc agcggcagaa gaagatgcag 1800  
gcagctgagt tgggttattc tgataagagt cagaggtaac tcccggtgc gtgtgtttaa 1860  
cggtggaggg cagttagtc tgagcagtagc tcgttgctgc cgccgcgc accagacata 1920  
atagctgaca gactaacaga ctgttcctt ccatgggtct tttctgcagt caccgtcgtc 1980  
gacctaagaa ttcaactcg agcaagtcta gaaaggcgcg ccaagatatac aaggatccac 2040  
tacgcgttag agctcgctga tcagcctcgatgtgccttc tagttgccag ccatctgttgc 2100  
tttgcgccttc ccccgctgcct tccttgaccc tggaagggtgc cactccact gtccttcct 2160  
aataaaatga ggaaattgca tcgcattgtc tgagtaggtg tcattctatt ctggggggtg 2220  
gggtggggca ggacagcaag ggggaggatt ggaaagacaa tagcaggcat gctggggagc 2280  
tcttcgcctt cctcgctcac tgactcgctg cgctcggtcg ttccggctgcg gcgagcggta 2340  
tcagctcact caaaggcggt aatacggtta tccacagaat cagggataa cgcaggaaag 2400  
aacatgtgag caaaaggcca gaaaaaggcc aggaaccgtaa aaaaggccgc gttgctggcg 2460  
ttttccata ggctccgcccc ccctgacgag catcacaaaa atcgacgctc aagtcaagg 2520  
tggcgaaacc cgacaggact ataaagatac caggggtttc cccctggaaag ctccctcg 2580  
cgctctcctg ttccgaccctt gccgcttacc ggataacctgt ccgccttct cccttcggga 2640  
agcgtggcgc tttctcaatg ctcacgctgt aggtatctca gttcgggtgtaa ggtcgttcgc 2700

00000000000000000000000000000000

tccaaagctgg gctgtgtgca cgaacccccc gttcagcccg accgctgcgc cttatccgg 2760  
aactatcgtc ttgagtccaa cccggtaaga cacgacttat cgccactggc agcagccact 2820  
ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtggtgg 2880  
cctaactacg gctacactag aaggacaga tttggtatct gcgctctgct gaagccagtt 2940  
accttcggaa aaagagttgg tagctttga tccggcaaac aaaccaccgc tggtagcggt 3000  
ggtttttttg tttgcaagca gcagattacg cgcaaaaaaa aaggatctca agaagatcct 3060  
ttgatctttt ctacggggtc tgacgctca taggaacgaaa actcacgtt aaggattttg 3120  
gtcatgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt 3180  
aaatcaatct aaagtatata ttagttaact tggctgaca gttaccaatg cttaatcagt 3240  
gaggcaccta tctcagcgat ctgtctattt cgttcatoca tagttgcctg actccccgtc 3300  
gtgttagataa ctacgataacg ggagggctta ccatctggcc ccagtgctgc aatgataaccg 3360  
cgagacccac gtcaccggc tccagattt tcagaataa accagccagc cggaagggcc 3420  
gagcgcagaa gtggcctgc aactttatcc gcctccatcc agtctattaa ttgttgcgg 3480  
gaagctagag taagtagttc gccagttaat agttgcgca acgttgcgc cattgctaca 3540  
ggcatcgtgg tgtcacgctc gtcgttttgtt atggcttcat tcagctccgg ttcccaacga 3600  
tcaaggcgag ttacatgatc ccccatgtt tgcaaaaaag cggttagctc cttcggtct 3660  
ccgatcgttg tcagaagtaa gttggccgca gtgtatcac tcatggttat ggcagcactg 3720  
cataattctc ttactgtcat gccatccgta agatgcttt ctgtgactgg tgagtaactca 3780  
accaagtcat tctgagaata gtgtatgcgg cgaccgagtt gctctgccc ggcgtcaata 3840  
cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg aaaacgttct 3900  
tcggggcgaa aactctcaag gatcttacgg ctgtttagat ccagttcgat gtaaccact 3960  
cgtgcaccca actgatcttc agcatcttt acttcacca gcgttctgg gtgagcaaaa 4020  
acaggaaggc aaaatgccgc aaaaaaggga ataaggcga cacggaaatg ttgaataactc 4080  
atactcttcc ttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga 4140  
tacatatttg aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga 4200  
aaagtgccac ctgacgtcta agaaaccatt attatcatga cattaaccta taaaaatagg 4260  
cgtatcacga ggccctttcg tc 4282

<210> 6  
<211> 6299

DNA 序列

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: pNS34a

<220>

<221> CDS

<222> (1990) .. (4047)

<400> 6

cgcgcgtttc ggtgatgacg gtgaaaacct ctgacacatg cagctccgg agacggcac 60  
agcttgtctg taagcgatg ccgggagcag acaagccgt cagggcgcgt cagcgggtgt 120  
tggcggtgt cggggctggc ttaactatgc ggcattcagag cagattgtac tgagagtca 180  
ccatatgaag cttttgcaa aagcctaggc ctccaaaaaa gcctcctcac tacttctgga 240  
atagctcaga ggccgaggcg gcctcggcct ctgcataaat aaaaaaaaaatt agtcagccat 300  
ggggcggaga atggcggaa ctggcgggg aggaaattat tggctattgg ccattgcata 360  
cgttgtatct atatcataat atgtacattt atattggctc atgtccaata tgaccgccat 420  
gttgacattt attattgact agttatataat agtaatcaat tacgggtca ttagttcata 480  
gccccatataat ggagttccgc gttacataac ttacggtaaa tggcccgccct ggctgaccgc 540  
ccaaacgaccc cggccattt acgtcaataa tgacgtatgt tcccatagta acgccaatag 600  
ggactttcca ttgacgtcaa tgggtggagt atttacggta aactgcccac ttggcagttac 660  
atcaagtgtt tcatatgcca agtccgcccc ctattgacgt caatgacggt aaatggcccg 720  
cctggcatta tgcccagtac atgaccttac gggactttcc tacttggcag tacatctacg 780  
tattagtcattt cgctattacc atggatgatgc ggtttggca gtacaccaat gggcgtggat 840  
agcggtttga ctcacgggaa tttccaagtc tccacccat tgacgtcaat gggagttgt 900  
tttggcacca aaatcaacgg gactttccaa aatgtcgtaa taacccggcc cgggtgacgc 960  
aaatgggcgg taggcgtgtt cgggtggagg tctatataag cagagctcggt ttagtgaacc 1020  
gtcagatcgc ctggagacgc catccacgtt gtttgaccc tcatagaaga caccggacc 1080  
gatccagcctt ccgcggccgg gaacggtgca ttggAACGCG gattcccggtt gccaagagtgt 1140  
acgtaagtac cgcctataga ctctataggc acacccctt ggctcttatg catgctatac 1200  
tggtttggc ttggggccta tacacccccc ctccttatgc tataggtgat ggtatagttt 1260  
agcctatagg tgtgggttat tgaccattat tgaccactcc cctattggtg acgataacttt 1320  
ccattactaa tccataacat ggctttgc cacaactatc tctattggct atatgccaat 1380

actctgcct tcagagactg acacggactc tgtatTTTA caggatgggg tccatttatt 1440  
atttacaat tcacatatac aacaacgccc tccccgtgc ccgcagttt tattaaacat 1500  
agcgtggat ctccgacatc tcgggtacgt gttccggaca tgggtcttc tccggtagcg 1560  
gcggagcttc cacatccgag ccctggtccc atccgtccag cggctcatgg tcgctcgca 1620  
gctccttgct cctaacagtg gaggccagac ttaggcacag cacaatgccc accaccacca 1680  
gtgtgccgca caaggccgtg gcggtagggt atgtgtctga aaatgagctc ggagattggg 1740  
ctcgcacctg gacgcagatg gaagacttaa ggcagcggca gaagaagatg caggcagctg 1800  
agttgtttaa ttctgataag agtcagaggt aactcccgtt gcggtgttgt taacgggttga 1860  
gggcagtgtt gtctgagcag tactcggtgc tgccgcgcgc gccaccagac ataatagctg 1920  
acagactaac agactgttcc tttccatggg tctttctgc agtcaccgtc gtcgaccaa 1980  
gaattcacc atg gcg ccc atc acg gcg tac gcc cag cag aca agg ggc ctc 2031  
Met Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu  
1 5 10  
ctt ggg tgc ata atc acc agc cta act ggc cgg gac aaa aac caa gtg 2079  
Leu Gly Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val  
15 20 25 30  
gag ggt gag gtc cag att gtg tca act gct gcc caa acc ttc ctg gca 2127  
Glu Gly Glu Val Gln Ile Val Ser Thr Ala Ala Gln Thr Phe Leu Ala  
35 40 45  
acg tgc atc aat ggg gtg tgc tgg act gtc tac cac ggg gcc gga acg 2175  
Thr Cys Ile Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Thr  
50 55 60  
agg acc atc gcg tca ccc aag ggt cct gtc atc cag atg tat acc aat 2223  
Arg Thr Ile Ala Ser Pro Lys Gly Pro Val Ile Gln Met Tyr Thr Asn  
65 70 75  
gta gac caa gac ctt gtg ggc tgg ccc gct tcg caa ggt acc cgc tca 2271  
Val Asp Gln Asp Leu Val Gly Trp Pro Ala Ser Gln Gly Thr Arg Ser  
80 85 90  
ttt aca ccc tgc act tgc ggc tcc tcg gac ctt tac ctg gtc acg agg 2319  
Leu Thr Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg  
95 100 105 110  
cac gcc gat gtc att ccc gtg cgc cgg cgg ggt gat agc agg ggc agc 2367  
His Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser  
115 120 125  
ctg ctg tcg ccc cgg ccc att tcc tac ttg aaa ggc tcc tcg ggg ggt 2415  
Leu Leu Ser Pro Arg Pro Ile Ser Tyr Leu Lys Gly Ser Ser Gly Gly  
130 135 140  
ccg ctg ttg tgc ccc gcg ggg cac gcc gtg ggc ata ttt agg gcc qcg 2463



DRAFT 2/22/00

370

375

380

gtc gca ttg ggc atc aat gcc gtg gcc tac tac	cgc ggt ctt gac gtg	3183
Val Ala Leu Gly Ile Asn Ala Val Ala Tyr Tyr	Arg Gly Leu Asp Val	
385	390	395
tcc gtc atc ccg acc agc ggc gat gtt gtc gtc	gtg gca acc gat gcc	3231
Ser Val Ile Pro Thr Ser Gly Asp Val Val Val	Ala Thr Asp Ala	
400	405	410
ctc atg acc ggc tat acc ggc gac ttc gac	tcg gtg ata gac tgc aat	3279
Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp	Ser Val Ile Asp Cys Asn	
415	420	425
430		
acg tgt gtc acc cag aca gtc gat ttc agc	ctt gac cct acc ttc acc	3327
Thr Cys Val Thr Gln Thr Val Asp Phe Ser	Leu Asp Pro Thr Phe Thr	
435	440	445
att gag aca atc acg ctc ccc caa gat gtc	tcc cgc act caa cgt	3375
Ile Glu Thr Ile Thr Leu Pro Gln Asp Ala Val	Ser Arg Thr Gln Arg	
450	455	460
cgg ggc agg act ggc agg ggg aag cca ggc atc tac	aga ttt gtg gca	3423
Arg Gly Arg Thr Gly Arg Gly Lys Pro Gly Ile	Tyr Arg Phe Val Ala	
465	470	475
ccg ggg gag cgc ccc tcc ggc atg ttc gac tcg	tcc gtc ctc tgt gag	3471
Pro Gly Glu Arg Pro Ser Gly Met Phe Asp Ser	Ser Val Leu Cys Glu	
480	485	490
tgc tat gac gca ggc tgt gct tgg tat gag ctc	acg ccc gcc gag act	3519
Cys Tyr Asp Ala Gly Cys Ala Trp Tyr Glu Leu	Thr Pro Ala Glu Thr	
495	500	510
aca gtt agg cta cga gcg tac atg aac acc ccg	ggg ctt ccc gtg tgc	3567
Thr Val Arg Leu Arg Ala Tyr Met Asn Thr Pro	Gly Leu Pro Val Cys	
515	520	525
cag gac cat ctt gaa ttt tgg gag ggc gtc ttt	aca ggc ctc act cat	3615
Gln Asp His Leu Glu Phe Trp Glu Gly Val Phe	Thr Gly Leu Thr His	
530	535	540
ata gat gcc cac ttt cta tcc caa aca aag cag	agt ggg gag aac ctt	3663
Ile Asp Ala His Phe Leu Ser Gln Thr Lys Gln	Ser Gly Glu Asn Leu	
545	550	555
cct tac ctg gta gcg tac caa gcc acc gtg tgc	gct agg gct caa gcc	3711
Pro Tyr Leu Val Ala Tyr Gln Ala Thr Val Cys	Ala Arg Ala Gln Ala	
560	565	570
cct ccc cca tcg tgg gac cag atg tgg aag tgc	att cgc ctc aag	3759
Pro Pro Pro Ser Trp Asp Gln Met Trp Lys Cys	Leu Ile Arg Leu Lys	
575	580	585
ccc acc ctc cat ggg cca aca ccc ctg cta tac	aga ctg ggc gct gtt	3807
Pro Thr Leu His Gly Pro Thr Pro Leu Leu Tyr	Arg Leu Gly Ala Val	
595	600	605



atgaagttt aaatcaatct aaagtatata tgagtaaact tggctcgaca gttaccaatg 5247  
cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca tagttgcctg 5307  
actccccgtc gtgtagataa ctacgatacg ggagggctta ccatctgcc ccagtgctgc 5367  
aatgataccg cgagacccac gctcaccggc tccagattt tcagcaataa accagccagc 5427  
cggaaggggcc gagcgcagaa gtggtcctgc aactttatcc gcctccatcc agtctattaa 5487  
ttgttgcgg gaagctagag taagtagtc gccagttaaat agttgcgcac acgttgttgc 5547  
cattgctaca ggcatcgtag tgtcacgctc gtcgttgat atggcttcat tcagctccgg 5607  
ttcccaacga tcaaggcgag ttacatgatc cccatgttg tgcaaaaaag cggttagctc 5667  
cttcggtcct ccgatcgtag tcagaagtaa gttggccgca gtgttatcac tcattgttat 5727  
ggcagcactg cataattctc ttactgtcat gccatcgta agatgcttt ctgtgactgg 5787  
tgagtactca accaagtcat tctgagaata gtgtatgcgg cgaccgagtt gctcttgc 5847  
ggcgtcaata cgggataata cgcgcacaca tagcagaact taaaagtgc tcattattgg 5907  
aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat 5967  
gtaacccact cgtgcaccca actgatctc agcatcttt actttcacca gcgttctgg 6027  
gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataaggcgac cacggaaatg 6087  
ttgaataactc atactttcc ttttcaata ttattgaagc atttatcagg gttattgtct 6147  
catgagcgga tacatatttg aatgtattt gaaaaataaa caaatagggg ttccgcgcac 6207  
atttccccga aaagtgcac ctgacgtcta agaaaccatt attatcatga cattaaccta 6267  
taaaaatagg cgtatcacga ggcctttcg tc 6299

<210> 7  
<211> 686  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: pNS34a

<400> 7  
Met Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu Leu Gly  
1 5 10 15  
  
Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu Gly  
20 25 30  
  
Glu Val Gln Ile Val Ser Thr Ala Ala Gln Thr Phe Leu Ala Thr Cys  
35 40 45

Ile Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Thr Arg Thr  
50 55 60

Ile Ala Ser Pro Lys Gly Pro Val Ile Gln Met Tyr Thr Asn Val Asp  
65 70 75 80

Gln Asp Leu Val Gly Trp Pro Ala Ser Gln Gly Thr Arg Ser Leu Thr  
85 90 95

Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His Ala  
100 105 110

Asp Val Ile Pro Val Arg Arg Gly Asp Ser Arg Gly Ser Leu Leu  
115 120 125

Ser Pro Arg Pro Ile Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro Leu  
130 135 140

Leu Cys Pro Ala Gly His Ala Val Gly Ile Phe Arg Ala Ala Val Cys  
145 150 155 160

Thr Arg Gly Val Ala Lys Ala Val Asp Phe Ile Pro Val Glu Asn Leu  
165 170 175

Glu Thr Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Ser Pro Pro  
180 185 190

Val Val Pro Gln Ser Phe Gln Val Ala His Leu His Ala Pro Thr Gly  
195 200 205

Ser Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly Tyr  
210 215 220

Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly  
225 230 235 240

Ala Tyr Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr Gly  
245 250 255

Val Arg Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly  
260 265 270

Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile  
275 280 285

Ile Cys Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile  
290 295 300

Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val Val  
305 310 315 320

Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn  
325 330 335

Ile Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr Gly  
340 345 350

Lys Ala Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu Ile Phe  
355 360 365

Cys His Ser Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Ala  
370 375 380

Leu Gly Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val  
385 390 395 400

Ile Pro Thr Ser Gly Asp Val Val Val Ala Thr Asp Ala Leu Met  
405 410 415

Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys  
420 425 430

Val Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu  
435 440 445

Thr Ile Thr Leu Pro Gln Asp Ala Val Ser Arg Thr Gln Arg Arg Gly  
450 455 460

Arg Thr Gly Arg Gly Lys Pro Gly Ile Tyr Arg Phe Val Ala Pro Gly  
465 470 475 480

Glu Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr  
485 490 495

Asp Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Thr Val  
500 505 510

Arg Leu Arg Ala Tyr Met Asn Thr Pro Gly Leu Pro Val Cys Gln Asp  
515 520 525

His Leu Glu Phe Trp Glu Gly Val Phe Thr Gly Leu Thr His Ile Asp  
530 535 540

Ala His Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Leu Pro Tyr  
545 550 555 560

Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro  
565 570 575

Pro Ser Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr  
580 585 590

Leu His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn  
595 600 605

Glu Ile Thr Leu Thr His Pro Val Thr Lys Tyr Ile Met Thr Cys Met  
610 615 620

Ser Ala Asp Leu Glu Val Val Thr Ser Thr Trp Val Leu Val Gly Gly  
625 630 635 640

Val Leu Ala Ala Leu Ala Ala Tyr Cys Leu Ser Thr Gly Cys Val Val  
645 650 655

Ile Val Gly Arg Val Val Leu Ser Gly Lys Pro Ala Ile Ile Pro Asp  
660 665 670

Arg Glu Val Leu Tyr Arg Glu Phe Asp Glu Met Glu Glu Cys  
675 680 685

<210> 8  
<211> 19912  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: pd.deltaNS3NS5

<220>  
<221> CDS  
<222> (12745)..(18057)

<400> 8  
atcgatccta ccccttgcgc taaagaagta tatgtgccta ctaacgcttg tctttgtctc 60  
tgtcaactaaa cactggatta ttactcccag atacttattt tggactaatt taaatgattt 120  
cgatatcaacg ttcttaatat cgctgaatct tccacaattt atgaaagtag cttaggaagag 180  
gaattggtat aaagttttt 5' tttttgtaaa tctcgaagta tactcaaacg aatttagtat 240  
tttctcagtg atctccaga tgcttcacc ctcacttaga agtgctttaa gcatttttt 300  
actgtggcta ttcccttat ctgcttcctc cgatgattcg aactgtaatt gcaaactact 360  
tacaatataca gtgatatacg attgatgttt ttgtccatag taaggaataa ttgtaaattc 420  
ccaaggcagga atcaatttct ttaatgaggc ttccagaatt gttgctttt gcgtcttgc 480  
tttaaactgg agtgatttat tgacaatatac gaaactcagc gaattgctta tgatagtatt 540  
atagctcatg aatgtggctc tcttgattgc tgccgtta tgtgtaatca tccaacataaa 600  
ataggttagt tcagcagcac ataatgctat tttctcacct gaaggtctt caaaccttc 660  
cacaaactga cgaacaagca ccttaggtgg tggtttacat aatataatcaa attgtggcat 720  
gcttagcgcc gatcttgtgt gcaattgata tctagttca actactctat ttatcttgc 780  
tcttgcagta ttcaaacacg ctaactcgaa aaactaactt taattgtcct gtttgtctcg 840  
cggtatcttc atttcatatt taaaaatgc acctttgctg cttttcctta attttagac 900  
ggccccgcagg ttctgtttgc ggtactatct tgtgataaaaa agttgtttt acatgtgatc 960  
tgcacagatt ttataatgta ataagcaaga atacattatc aaacgaacaa tactggtaaa 1020  
agaaaaaccaa aatggacgac attgaaacag ccaagaatct gacggtaaaa gcacgtacag 1080  
1140

002030405060708090

cttatagcgt ctggatgta tgcggctgt ttattgaaat gattgctcct gatgttagata 1200  
ttgatataga gagtaaacgt aagtctgatg agctactctt tccaggatat gtcataaggc 1260  
ccatggaatc tctcacaacc gtagggcgt atggcttga ttctagcgca gaagattcca 1320  
gcgtatctc tgactccagt gctgaggtaa tttgcctgc tgcgaagatg gttaaggaaa 1380  
ggtttgattc gattgaaat ggtatgctt cttcacaaga agcaagtcag gctgccatag 1440  
atttgatgct acagaataac aagctgttag acaatagaaa gcaactatac aaatctattg 1500  
ctataataat aggaagattt cccgagaaag acaagaagag agctaccgaa atgctcatga 1560  
aaaaaatgga ttgtacacag ttattagtcc caccagctcc aacggaagaa gatgttatga 1620  
agctcgtaag cgtcgttacc caattgctt ctttagttcc accagatcgt caagctgctt 1680  
taataggtga ttatttcattc ccggaatctc taaaggatat attcaatagt ttcaatgaac 1740  
tggcggcaga gaatcgaaa cagcaaaaaa agagttagtt ggaaggaagg actgaagtga 1800  
accatgctaa tacaaatgaa gaagttccct ccagcgaaac aagaagtaga gacacaaatg 1860  
caagaggagc atataaatta caaaacacca tcactgaggg ccctaaagcg gttcccacga 1920  
aaaaaaggag agtagcaacg agggtaaggg gtagaaaatc acgtaatact tctagggtat 1980  
gatccaatat caaaggaaat gatagcattt aaggatgaga ctaatccat tgaggagtgg 2040  
cagcatatac aacagctaaa ggtagtgtt gaaaggaaagca tacgatacccc cgcatggaat 2100  
gggataatat cacaggaggt actagactac ctttcatttcc acataaatac acgcatataa 2160  
gtacgcattt aagcataaac acgcactatg ccgttcttcatgtatata tatatacagg 2220  
caacacgcag atatagggtt gacgtgaaca gtgagctgtt tgtgcgcagc tcgcgttgca 2280  
tttcggaaag cgctcgaaaa cggaaacgctt ttgaagttcc tattccgaaat ttcctatttt 2340  
ctagaaagta taggaacttc agagcgctt tgaaaaccaa aagcgctctg aagacgcact 2400  
ttcaaaaaac caaaaacgca ccggactgtt acgagctact aaaatattgc gaataccgct 2460  
tccacaaaca ttgctcaaaa gatatcttt gctatataatc tctgtgttat atccctatata 2520  
aacctaccca tccacatccc gtccttggaa cttgcatttta aactcgacct ctacatcaac 2580  
aggcttccaa tgcttttcaaa attttactgtt caagtagacc catacggtt taatatgtt 2640  
ctcttcataa tgtaagcttta tctttatcgaa atcggtgtt aaactactac cgcgataaac 2700  
ctttacgggtt ccctgagattt gatattgttcc cttagtata tgatataaaga cacttttggaa 2760  
ctttgtacga cgaattttga gtttgcctt cctctggctt tttccatata tcctgtcgcc 2820  
gaaacatgtt gcttaaaact ccaagcggtt ggagaccgtt aaaggttaat aggacagccg 2880

0 1 2 3 4 5 6 7 8 9

tattatctcc gcctcagttt gatcttccgc ttcagactgc ~~cattttcac~~ ataatgaatc 2940  
tatttcaccc cacaatcctt catccgcctc cgcatcttgt ~~tccgttaaac~~ tattgacttc 3000  
atgttgtaca ttgttttagtt cacgagaagg gtccttca ~~ggcggtagct~~ cctgatctcc 3060  
tatatgacct ttatcctgtt ctcttccac aaacttagaa atgtattcat gaattatgga 3120  
gcacctaata acattttca aggccggagaa gtttggcca gatgccaaat atgcttgaca 3180  
tgaaaacgtg agaatgaatt tagtattatt gtgatattct gaggcaattt tattataatc 3240  
tcgaagataa gagaagaatg cagtgacctt tgtattgaca aatggagatt ccatgtatct 3300  
aaaaaaatacg cctttaggcc ttctgatacc ctccccctg cggttagcg tgcctttac 3360  
attaatatct aaaccctctc cgatggtggc cttaactga ctaataaaatg caaccgatat 3420  
aaactgtgat aattctgggt gatttatgat tcgatcgaca attgtattgt acactagtgc 3480  
aggatcaggc caatccagtt cttttcaat taccgggttg tcgtctgtat tcagtagatg 3540  
tccaacaaat gcaaatgcta acgtttgtt ttttttataa ttgtcaggaa ctggaaaagt 3600  
cccccttgtc gtctcgatta cacacctact ttcatcgta accatagggtt ggaagtgctg 3660  
cataatacat tgcttaatac aagcaagcag tccctcgcca ttcatatccc agttatccc 3720  
cattacagct gatgtcattt tatatcagcg ctgtaaaaat ctatctgtt cagaaggttt 3780  
tcgcggtttt tataaacaaa actttcgta cggaaatcgag caatcacccc agctgcgtat 3840  
ttggaaattt cggaaaaagt agagcaacgc gagttgcatt ttttacacca taatgcata 3900  
ttaacttcga gaagggatta aggctaattt cactagtatg tttcaaaaac ctcaatctgt 3960  
ccattgaatg ctttataaaa cagctataga ttgcataaaaa gagtttagcta ctcaatgctt 4020  
tttgtcaaaat cttactgtatg atgatgtgtc tacttcagg cgggtctgtt gtaaggagaa 4080  
tgacattata aagctggcac tttagaattcc acggactata gactatacta gtataactccg 4140  
tctactgtac gatacacttc cgctcagggtc cttgtccctt aacgaggcct taccactctt 4200  
ttgttactct attgatccag ctcagcaaaag gcagtgtat ctaagattct atcttcgcga 4260  
tgttagaaaaa ctagctagac cgagaaaagag actagaaaatg caaaaggcac ttctacaatg 4320  
gctgccatca ttattatccg atgtgacgct gcattttttt tttttttttt tttttttttt 4380  
tttttttttt tttttttttt tttttttgtt caaatatcat aaaaaaaagag aatctttta 4440  
agcaaggatt ttcttaactt ctccggcgcac agcatcaccg acttcgggtgg tactgttgaa 4500  
accaccaaaa tcaccagttc tgataactgc atccaaaacc ttttttaactg catcttcaat 4560  
ggctttaccc ttctcaggca agttcaatga caatttcaac atcattgcag cagacaagat 4620

DNA sequence analysis

agtggcgata gggttgacct tattctttgg caaatctgga gcggaaaccat ggcatggttc 4680  
gtacaaaacca aatgcggtgt tcttgtctgg caaagaggcc aaggacgcag atggcaacaa 4740  
acccaaggag cctgggataa cggaggcttc atcggagatg atatcaccaa acatgttgct 4800  
ggtgattata ataccattta ggtgggttgg gttcttaact aggatcatgg cggcagaatc 4860  
aatcaattga tgttgaactt tcaatgttagg gaattcggttc ttgatggttt cctccacagt 4920  
tttctccat aatcttgaag aggccaaaac attagcttta tccaaggacc aaataggcaa 4980  
tggtggctca tggttaggg ccatgaaagc ggccattctt gtgattctt gcacttctgg 5040  
aacggtgtat tgttcactat cccaagcgc accatcacca tcgtcttcct ttctcttacc 5100  
aaagtaaata cctccacta attctctaac aacaacgaag tcagtagctt tagcaaattg 5160  
tggcttgatt ggagataagt ctaaaagaga gtcggatgca aagttacatg gtcttaagtt 5220  
ggcgtacaat tgaagttctt tacggatttt tagtaaacct tgttcaggc taacactacc 5280  
ggtaccccat ttaggaccac ccacagcacc taacaaaacg gcatcaggct tcttggaggc 5340  
ttccagcgcc tcatctggaa gtggaacacc tgttagcatcg atagcagcac caccaattaa 5400  
atgattttcg aaatcgaact tgacattgga acgaacatca gaaatagctt taagaacctt 5460  
aatggcttcg gctgtgattt cttgaccaac gtggcacct ggcaaaacga cgatcttctt 5520  
aggggcagac attacaatgg tatatcctt aaatatataa aaaaaaaaaa aaaaaaaaaa 5580  
aaaaaaaaaa atgcagcttc tcaatgatat tcgaatacgc tttgaggaga tacagcctaa 5640  
tatccgacaa actgtttac agattacga tcgtacttgt tacccatcat tgaattttga 5700  
acatccgaac ctgggagttt tccctgaaac agatagtata tttgaacctg tataataata 5760  
tatagtctag cgcttacgg aagacaatgt atgtatttcg gttcctggag aaactattgc 5820  
atctattgca tagtaatct tgcacgtcgc atccccgtt cattttctgc gtttccatct 5880  
tgcaactcaa tagcatatct ttgttaacga agcatctgtg cttcattttg tagaacaaaaa 5940  
atgcaacgcg agagcgctaa ttttcaaac aaagaatctg agctgcattt ttacagaaca 6000  
gaaatgcaac gcgaaagcgc tattttacca acgaagaatc tgtgcttcattttgtaaaa 6060  
caaaaaatgca acgcgagagc gctaattttt caaacaaga atctgagctg catttttaca 6120  
gaacagaaaat gcaacgcgag agcgctattt taccacaaa gaatctatac ttcttttttg 6180  
ttctacaaaaa atgcatcccg agagcgctat ttttctaaca aagcatcttta gattactttt 6240  
tttctccctt gtgcgctcta taatgcagtc tcttgataac tttttgcact gtaggtccgt 6300  
taaggtaga agaaggctac tttgggtgtct atttctctt ccataaaaaa agcctgactc 6360

cacttccgc gttactgat tactagcgaa gctgcgggtg cattttca agataaaggc 6420  
atccccgatt atattctata ccgatgtgga ttgcgcatac ttgtgaaca gaaagtgata 6480  
gcgttcatgtatc ttcttcattt gtcagaaaat tatgaacggg ttcttctatt ttgtctctat 6540  
atactacgta taggaaatgt ttacatttc gtattgtttt cgattcaactc tatgaatagt 6600  
tcttactaca attttttgtt ctaaagagta atactagaga taaacataaa aatgttagag 6660  
gtcgagtttta gatgcaagtt caaggagcga aaggtggatg ggttagttat atagggatat 6720  
agcacagaga tatatacgaa agagatactt ttgagcaatg tttgtgaaag cggtattcgc 6780  
aatatTTTtag tagctcgta cagtcgggtg cgTTTTGGT ttttggaaag tgcgtttca 6840  
gagcgctttt ggTTTcaaa agcgctctga agttccatata ctTTCTAGAG aataggaact 6900  
tcggaatagg aacttcaaag cgTTTCCGAA aacgAGCGCT tccggaaatg caacgcgagc 6960  
tgcgcacata cagctcaactg ttcacgtcgc acctatatct gcgtgttgcc tgtatata 7020  
tatacatgag aagaacggca tagtgcgtgt ttatgcctaa atgcgtactt atatgcgtct 7080  
atttatgttag gatgaaaggt agtctagttac ctccctgtat attatcccattccatgcggg 7140  
gtatcgtagt cttccTTCAg cactaccTTT tagctgttct atatgctgcc actcctcaat 7200  
tggattagtc tcATCCTTCA atgctatcat ttcctttgat attggatcat atgcata 7260  
ccgagaaact agtgcgaagt agtgatcagg tattgctgtt atctgatgag tatacgTTT 7320  
cctggccacg gcagaagcac gcttATCCTG ccaatttccc acaacattag tcaactccgt 7380  
tagGCCCTTC attgaaagaa atgaggtcat caaatgtctt ccaatgtgag attttggggcc 7440  
atttttata gcaaagattt aataaggcgc atttttcttc aaagctttat tgtacgatct 7500  
gactaaggta tcttttaata attggatttc ctgtttattt cttgaagaat tgccggcct 7560  
atttactcgt tttaggactg gttcagaatt cctcaaaaat tcattccaaat atacaagtgg 7620  
atcgatgata agctgtcaaa catgagaatt ctgtaaagacg aaagggcctc gtgatacgc 7680  
tattttata ggttaatgtc atgataataa tggTTTCTTA gacgtcagggt ggcactttc 7740  
ggggaaatgt gcgcggaaacc cctattgtt tattttctta aatacattca aatatgtatc 7800  
cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg aagagtatga 7860  
gtattcaaca ttccgtgtc gcccTTTATTc cctttttgc ggcattttgc ttccctgtt 7920  
ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag 7980  
tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagttt cggcccgaaag 8040  
aacgtttcc aatgatgagc aTTTTAAAG ttctgctatg tggcgcggta ttatcccgtg 8100

ttgacgcccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg 8160  
agtactcacc agtcacagaa aagcatctta cggtggcat gacagtaaga gaattatgca 8220  
gtgctgccat aaccatgagt gataaacactg cggccaactt acttctgaca acgatcgag 8280  
gaccgaagga gctaaccgct ttttgacaca acatggggga tcatgtact cgccttgatc 8340  
gttggaaacc ggagctgaat gaagccatac caaaacgacga gcgtgacacc acgatgcctg 8400  
cagcaatggc aacaacgttg cgcaaactat taactggcga actacttact cttagttccc 8460  
ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg 8520  
cccttccggc tggctggttt attgctgata aatctggagc cggtagcgt gggtctcgcg 8580  
gtatcattgc agcactgggg ccagatggta agccctcccg tatcgttagtt atctacacga 8640  
cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac 8700  
tgattaagca ttggtaactg tcagaccaag ttactcata tatacttttag attgatttaa 8760  
aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgcataat ctcatgacca 8820  
aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag 8880  
gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaaca aaaaaaccac 8940  
cgctaccagc ggtgggttgt ttgccggatc aagagctacc aactctttt ccgaaggtaa 9000  
ctggcttcag cagagcgcag ataccaaata ctgtccttct agttagccg tagttaggcc 9060  
accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaata ctgttaccag 9120  
tggctgctgc cagtggcgat aagtcgtgc ttaccgggtt ggactcaaga cgatagttac 9180  
cggataaggc gcagcggcgtg ggctgaacgg ggggttcgtg cacacagccc agttggagc 9240  
gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc 9300  
ccgaaggagaa aaaggcggac aggtatccgg taagcggcag ggtcggaaca ggagagcgca 9360  
cgagggagct tccaggggaa aacgcctgtt atctttatag tcctgtcggg tttcgccacc 9420  
tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg 9480  
ccagcaacgc ggcctttta cggttcctgg cctttgctg gcctttgct cacatgttct 9540  
ttcctgcgtt atccccgtat tctgtggata accgtattac cgcctttgag tgagctgata 9600  
ccgctcgccg cagccgaacg accgagcgcga gcgagtcagt gagcgaggaa gcggaaagagc 9660  
gcctgatgcgtgtat cttacgcata tgcgtggat ttcacacccgc atatggtgca 9720  
ctctcagttac aatctgctct gatgcgcata agttaagccaa gtatacactc cgctatcgct 9780  
acgtgactgg gtcatggctg cgcccccgcaca cccgccaaca cccgctgacg cgccctgacg 9840



B222222222222222222222222

caataaccagt ccggatcaac tggcaccatc tctcccgtag tcccatctaa ttttcttcc 11640  
 ggatgagggtt ccagatatac cgcaacacctt ttattatggt tccctgagg gaataataga 11700  
 atgtcccatt cgaaatcacc aattctaaac ctgggcgaat tgtatccgg gtttgtaac 11760  
 tcgttccagt caggaatgtt ccacgtgaag ctatcttcca gcaaagtctc cacttcttca 11820  
 tcaaattgtg gagaatactc ccaatgctct tatctatggg acttccggga aacacagtc 11880  
 cgatacttcc caattcgtct tcagagctca ttgtttgtt gaagagacta atcaaagaat 11940  
 cgtttctca aaaaaattaa tatcttaact gatagttga tcaaaggggc aaaacgtagg 12000  
 ggcaaacaaa cgaaaaatc gtttctcaa ttttctgatg ccaagaactc taaccagtct 12060  
 tatctaaaaa ttgccttatg atccgtctct ccgggttadag cctgtgtaac tgattaatcc 12120  
 tgcccttcta atcaccattc taatgttttta attaaggat tttgtcttca ttaacggctt 12180  
 tcgctcataa aaatgttatg acgtttgcc cgccaggccggg aaaccatcca cttcacgaga 12240  
 ctgatctcct ctgccccaaac accgggcattc tccaaacttat aagttggaga aataagagaa 12300  
 tttcagattt agagaatgaa aaaaaaaaaac ctttagttca taggtccatt ctcttagcgc 12360  
 aactacagag aacaggggca caaacaggca aaaaacgggc acaacctcaa tggagtgtatg 12420  
 caacctgcct ggagtaaatg atgacacaag gcaattgacc cacgtatgtatc tctatctcat 12480  
 tttcttacac cttctattac cttctgtct ctgtgattt gaaaaagctg aaaaaaaaaagg 12540  
 ttgaaaccag ttccctgaaa ttattccctt acttgactaa taagtatata aagacggtag 12600  
 gtattgattt taattctgtt aatctatttc taaacttct taaattctac ttttatagtt 12660  
 agtctttttt ttagtttaa aacaccaaga acttagtttca gaataaacac acataaaca 12720  
 acaagcttac aaaacaaatt cacc atg gct gca tat gca gct cag ggc tat 12771  
 Met Ala Ala Tyr Ala Ala Gln Gly Tyr  
 1 5  
 aag gtg cta gta ctc aac ccc tct gtt gct gca aca ctg ggc ttt ggt 12819  
 Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly  
 10 15 20 25  
 gct tac atg tcc aag gct cat ggg atc gat cct aac atc agg acc ggg 12867  
 Ala Tyr Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr Gly  
 30 35 40  
 gtg aga aca att acc act ggc agc ccc atc acg tac tcc acc tac ggc 12915  
 Val Arg Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly  
 45 50 55  
 aag ttc ctt gcc gac ggc ggg tgc tcg ggg ggc gct tat gac ata ata 12963  
 Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile  
 60 65 70

att tgt gac gag tgc cac tcc acg gat gcc aca tcc atc ttg ggc att Ile Cys Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile	75                   80                   85	13011
ggc act gtc ctt gac caa gca gag act gcg ggg gcg aga ctg gtt gtg Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val Val	90                   95                   100                   105	13059
ctc gcc acc gcc acc cct ccg ggc tcc gtc act gtg ccc cat ccc aac Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn	110                   115                   120	13107
atc gag gag gtt gct tcc acc acc gga gag atc cct ttt tac ggc Ile Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr Gly	125                   130                   135	13155
aag gct atc ccc ctc gaa gta atc aag ggg ggg aga cat ctc atc ttc Lys Ala Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu Ile Phe	140                   145                   150	13203
tgt cat tca aag aag aag tgc gac gaa ctc gcc gca aag ctg gtc gca Cys His Ser Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Ala	155                   160                   165	13251
ttg ggc atc aat gcc gtc tac tac cgc ggt ctt gac gtg tcc gtc Leu Gly Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val	170                   175                   180                   185	13299
atc ccg acc agc ggc gat gtt gtc gtc gtg gca acc gat gcc ctc atg Ile Pro Thr Ser Gly Asp Val Val Val Ala Thr Asp Ala Leu Met	190                   195                   200	13347
acc ggc tat acc ggc gac ttc gac tcg gtg ata gac tgc aat acg tgt Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys	205                   210                   215	13395
gtc acc cag aca gtc gat ttc agc ctt gac cct acc ttc acc att gag Val Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu	220                   225                   230	13443
aca atc acg ctc ccc caa gat gct gtc tcc cgc act caa cgt cgg ggc Thr Ile Thr Leu Pro Gln Asp Ala Val Ser Arg Thr Gln Arg Arg Gly	235                   240                   245	13491
agg act ggc agg ggg aag cca ggc atc tac aga ttt gtg gca ccg ggg Arg Thr Gly Arg Gly Lys Pro Gly Ile Tyr Arg Phe Val Ala Pro Gly	250                   255                   260                   265	13539
gag cgc ccc tcc ggc atg ttc gac tcg tcc gtc ctc tgt gag tgc tat Glu Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr	270                   275                   280	13587
gac gca ggc tgt gct tgg tat gag ctc acg ccc gcc gag act aca gtt Asp Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Thr Val	285                   290                   295	13635
agg cta cga gcg tac atg aac acc ccg ggg ctt ccc gtg tgc cag gac		13683

Arg	Leu	Arg	Ala	Tyr	Met	Asn	Thr	Pro	Gly	Leu	Pro	Val	Cys	Gln	Asp				
300																310			
cat	ctt	gaa	ttt	tgg	gag	ggc	gtc	ttt	aca	ggc	ctc	act	cat	ata	gat		13731		
His	Leu	Glu	Phe	Trp	Glu	Gly	Val	Phe	Thr	Gly	Leu	Thr	His	Ile	Asp				
315																325			
gcc	cac	ttt	cta	tcc	cag	aca	aag	cag	agt	ggg	gag	aac	ctt	cct	tac		13779		
Ala	His	Phe	Leu	Ser	Gln	Thr	Lys	Gln	Ser	Gly	Glu	Asn	Leu	Pro	Tyr				
330																340	345		
ctg	gta	gcg	tac	caa	gcc	acc	gtg	tgc	gct	agg	gct	caa	gcc	cct	ccc		13827		
Leu	Val	Ala	Tyr	Gln	Ala	Thr	Val	Cys	Ala	Arg	Ala	Gln	Ala	Pro	Pro				
350																355	360		
cca	tcg	tgg	gac	cag	atg	tgg	aag	tgt	ttg	att	cgc	ctc	aag	ccc	acc		13875		
Pro	Ser	Trp	Asp	Gln	Met	Trp	Lys	Cys	Leu	Ile	Arg	Leu	Lys	Pro	Thr				
365																370	375		
ctc	cat	ggg	cca	aca	ccc	ctg	cta	tac	aga	ctg	ggc	gct	gtt	cag	aat		13923		
Leu	His	Gly	Pro	Thr	Pro	Leu	Leu	Tyr	Arg	Leu	Gly	Ala	Val	Gln	Asn				
380																385	390		
gaa	atc	acc	ctg	acg	cac	cca	gtc	acc	aaa	tac	atc	atg	aca	tgc	atg		13971		
Glu	Ile	Thr	Leu	Thr	His	Pro	Val	Thr	Lys	Tyr	Ile	Met	Thr	Cys	Met				
395																400	405		
tcg	gcc	gac	ctg	gag	gtc	gtc	acg	agc	acc	tgg	gtg	ctc	gtt	ggc	ggc		14019		
Ser	Ala	Asp	Leu	Glu	Val	Val	Thr	Ser	Thr	Trp	Val	Leu	Val	Gly	Gly				
410																415	420	425	
gtc	ctg	gct	ttg	gcc	gcg	tat	tgc	ctg	tca	aca	ggc	tgc	gtg	gtc			14067		
Val	Leu	Ala	Ala	Leu	Ala	Ala	Tyr	Cys	Leu	Ser	Thr	Gly	Cys	Val	Val				
																430	435	440	
ata	gtg	ggc	agg	gtc	gtc	ttg	tcc	ggg	aag	ccg	gca	atc	ata	cct	gac		14115		
Ile	Val	Gly	Arg	Val	Val	Leu	Ser	Gly	Lys	Pro	Ala	Ile	Ile	Pro	Asp				
																445	450	455	
agg	gaa	gtc	ctc	tac	cga	gag	ttc	gat	gag	atg	gaa	gag	tgc	tct	cag		14163		
Arg	Glu	Val	Leu	Tyr	Arg	Glu	Phe	Asp	Glu	Met	Glu	Glu	Cys	Ser	Gln				
																460	465	470	
cac	tta	ccg	tac	atc	gag	caa	ggd	atg	atg	ctc	gcc	gag	ttc	aag			14211		
His	Leu	Pro	Tyr	Ile	Glu	Gln	Gly	Met	Met	Leu	Ala	Glu	Gln	Phe	Lys				
																475	480	485	
cag	aag	gcc	ctc	ggc	ctc	ctg	cag	acc	gcf	tcc	cgt	cag	gca	gag	gtt		14259		
Gln	Lys	Ala	Leu	Gly	Leu	Leu	Gln	Thr	Ala	Ser	Arg	Gln	Ala	Glu	Val				
																490	495	500	505
atc	gcc	cct	gct	gtc	cag	acc	aac	tgg	caa	aaa	ctc	gag	acc	ttc	tgg		14307		
Ile	Ala	Pro	Ala	Val	Gln	Thr	Asn	Trp	Gln	Lys	Leu	Glu	Thr	Phe	Trp				
																510	515	520	
gcg	aag	cat	atg	tgg	aac	ttc	atc	agt	ggg	ata	caa	tac	ttg	gcf	ggc		14355		
Ala	Lys	His	Met	Trp	Asn	Phe	Ile	Ser	Gly	Ile	Gln	Tyr	Leu	Ala	Gly				

DRAFT

	525	530	535	
ttg tca acg ctg cct ggt aac ccc gcc att gct tca ttg atg gct ttt				14403
Leu Ser Thr Leu Pro Gly Asn Pro Ala Ile Ala Ser Leu Met Ala Phe				
540	545	550		
aca gct gct gtc acc agc cca cta acc act agc caa acc ctc ctc ttc				14451
Thr Ala Ala Val Thr Ser Pro Leu Thr Ser Gln Thr Leu Leu Phe				
555	560	565		
aac ata ttg ggg ggg tgg gtg gct gcc cag ctc gcc gcc ccc ggt gcc				14499
Asn Ile Leu Gly Gly Trp Val Ala Ala Gln Leu Ala Ala Pro Gly Ala				
570	575	580	585	
gct act gcc ttt gtg ggc gct ggc tta gct ggc gcc gcc atc ggc agt				14547
Ala Thr Ala Phe Val Gly Ala Gly Leu Ala Gly Ala Ala Ile Gly Ser				
590	595	600		
gtt gga ctg ggg aag gtc ctc ata gac atc ctt gca ggg tat ggc gcg				14595
Val Gly Leu Gly Lys Val Leu Ile Asp Ile Leu Ala Gly Tyr Gly Ala				
605	610	615		
ggc gtg gcg gga gct ctt gtg gca ttc aag atc atg agc ggt gag gtc				14643
Gly Val Ala Gly Ala Leu Val Ala Phe Lys Ile Met Ser Gly Glu Val				
620	625	630		
ccc tcc acg gag gac ctg gtc aat cta ctg ccc gcc atc ctc tcg ccc				14691
Pro Ser Thr Glu Asp Leu Val Asn Leu Leu Pro Ala Ile Leu Ser Pro				
635	640	645		
gga gcc ctc gta gtc ggc gtg gtc tgt gca gca ata ctg cgc cgg cac				14739
Gly Ala Leu Val Val Gly Val Val Cys Ala Ala Ile Leu Arg Arg His				
650	655	660	665	
gtt ggc ccg ggc gag ggg gca gtg cag tgg atg aac cgg ctg ata gcc				14787
Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile Ala				
670	675	680		
ttc gcc tcc cgg ggg aac cat gtt tcc ccc acg cac tac gtg ccg gag				14835
Phe Ala Ser Arg Gly Asn His Val Ser Pro Thr His Tyr Val Pro Glu				
685	690	695		
agc gat gca gct gcc cgc gtc act gcc ata ctc agc agc ctc act gta				14883
Ser Asp Ala Ala Ala Arg Val Thr Ala Ile Leu Ser Ser Leu Thr Val				
700	705	710		
acc cag ctc ctg agg cga ctg cac cag tgg ata agc tcg gag tgt acc				14931
Thr Gln Leu Leu Arg Arg Leu His Gln Trp Ile Ser Ser Glu Cys Thr				
715	720	725		
act cca tgc tcc ggt tcc tgg cta agg gac atc tgg gac tgg ata tgc				14979
Thr Pro Cys Ser Gly Ser Trp Ile Arg Asp Ile Trp Asp Trp Ile Cys				
730	735	740	745	
gag gtg ttg agc gac ttt aag acc tgg cta aaa gct aag ctc atg cca				15027
Glu Val Leu Ser Asp Phe Lys Thr Trp Leu Lys Ala Lys Leu Met Pro				
750	755	760		

DNA SEQUENCING

cag ctg cct ggg atc ccc ttt gtg tcc tgc cag cgc	ggg tat aag ggg	15075
Gln Leu Pro Gly Ile Pro Phe Val Ser Cys Gln Arg	Gly Tyr Lys Gly	
765	770	775
gtc tgg cga ggg gac ggc atc atg cac act cgc tgc	cac tgt gga gct	15123
Val Trp Arg Gly Asp Gly Ile Met His Thr Arg	Cys His Cys Gly Ala	
780	785	790
gag atc act gga cat gtc aaa aac ggg acg atg	agg atc gtc ggt cct	15171
Glu Ile Thr Gly His Val Lys Asn Gly Thr Met	Arg Ile Val Gly Pro	
795	800	805
agg acc tgc agg aac atg tgg agt ggg acc ttc	ccc att aat gcc tac	15219
Arg Thr Cys Arg Asn Met Trp Ser Gly Thr	Phe Pro Ile Asn Ala Tyr	
810	815	820
acc acg ggc ccc tgt acc ccc ctt cct gcg	ccg aac tac acg ttc gcg	15267
Thr Thr Gly Pro Cys Thr Pro Leu Pro Ala	Pro Asn Tyr Thr Phe Ala	
830	835	840
cta tgg agg gtg tct gca gag gaa tac gtg	gag ata agg cag gtg ggg	15315
Leu Trp Arg Val Ser Ala Glu Glu Tyr Val	Glu Ile Arg Gln Val Gly	
845	850	855
gac ttc cac tac gtg acg ggt atg act act	gac aat ctt aaa tgc ccg	15363
Asp Phe His Tyr Val Thr Gly Met Thr	Asp Asn Leu Lys Cys Pro	
860	865	870
tgc cag gtc cca tcg ccc gaa ttt ttc aca gaa	ttg gac ggg gtg cgc	15411
Cys Gln Val Pro Ser Pro Glu Phe Phe Thr	Glu Leu Asp Gly Val Arg	
875	880	885
cta cat agg ttt gcg ccc ccc tgc aag	ccc ttg ctg cgg gag gag gta	15459
Leu His Arg Phe Ala Pro Pro Cys Lys	Pro Leu Leu Arg Glu Glu Val	
890	895	900
905		
tca ttc aga gta gga ctc cac gaa tac ccg	gta ggg tcg caa tta cct	15507
Ser Phe Arg Val Gly Leu His Glu Tyr Pro	Val Gly Ser Gln Leu Pro	
910	915	920
tgc gag ccc gaa ccg gac gtg gcc	gtg ttg acg tcc atg ctc act gat	15555
Cys Glu Pro Glu Pro Asp Val Ala	Val Leu Thr Ser Met Leu Thr Asp	
925	930	935
ccc tcc cat ata aca gca gag gca	gcc ggg cga agg ttg gcg agg gga	15603
Pro Ser His Ile Thr Ala Glu Ala	Ala Gly Arg Arg Leu Ala Arg Gly	
940	945	950
tca ccc ccc tct gtg gcc agc tcc tcg	gct agc cag cta tcc gct cca	15651
Ser Pro Pro Ser Val Ala Ser Ser Ala Ser	Gln Leu Ser Ala Pro	
955	960	965
tct ctc aag gca act tgc acc	gct aac cat gac tcc cct gat gct gag	15699
Ser Leu Lys Ala Thr Cys Thr Ala Asn His	Asp Ser Pro Asp Ala Glu	
970	975	980
985		
ctc ata gag gcc aac ctc cta tgg agg cag gag	atg ggc ggc aac atc	15747

Leu Ile Glu Ala Asn Leu Leu Trp Arg Gln Glu Met Gly Gly Asn Ile	990	995	1000	
acc agg gtt gag tca gaa aac aaa gtg gtg att ctg gac tcc ttc gat				15795
Thr Arg Val Glu Ser Glu Asn Lys Val Val Ile Leu Asp Ser Phe Asp	1005	1010	1015	
ccg ctt gtg gcg gag gag gac gag cgg gag atc tcc gta ccc gca gaa				15843
Pro Leu Val Ala Glu Glu Asp Glu Arg Glu Ile Ser Val Pro Ala Glu	1020	1025	1030	
atc ctg cgg aag tct cgg aga ttc gcc cag gcc ctg ccc gtt tgg gcg				15891
Ile Leu Arg Lys Ser Arg Arg Phe Ala Gln Ala Leu Pro Val Trp Ala	1035	1040	1045	
ccg ccg gac tat aac ccc cgg cta gtg gag acg tgg aaa aag ccc gac				15939
Arg Pro Asp Tyr Asn Pro Pro Leu Val Glu Thr Trp Lys Lys Pro Asp	1050	1055	1060	1065
tac gaa cca cct gtg gtc cat ggc tgc cgg ctt cca cct cca aag tcc				15987
Tyr Glu Pro Pro Val Val His Gly Cys Pro Leu Pro Pro Pro Lys Ser	1070	1075	1080	
cct cct gtg cct cgg cct cgg aag aag cgg acg gtg gtc ctc act gaa				16035
Pro Pro Val Pro Pro Pro Arg Lys Lys Arg Thr Val Val Leu Thr Glu	1085	1090	1095	
tca acc cta tct act gcc ttg gcc gag ctc gcc acc aga agc ttt ggc				16083
Ser Thr Leu Ser Thr Ala Leu Ala Glu Leu Ala Thr Arg Ser Phe Gly	1100	1105	1110	
agc tcc tca act tcc ggc att acg ggc gac aat acg aca aca tcc tct				16131
Ser Ser Ser Thr Ser Gly Ile Thr Gly Asp Asn Thr Thr Ser Ser	1115	1120	1125	
gag ccc gcc cct tct ggc tgc ccc ccc gac tcc gac gct gag tcc tat				16179
Glu Pro Ala Pro Ser Gly Cys Pro Pro Asp Ser Asp Ala Glu Ser Tyr	1130	1135	1140	1145
tcc tcc atg ccc ccc ctg gag ggg gag cct ggg gat ccg gat ctt agc				16227
Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu Ser	1150	1155	1160	
gac ggg tca tgg tca acg gtc agt agt gag gcc aac gcg gag gat gtc				16275
Asp Gly Ser Trp Ser Thr Val Ser Ser Glu Ala Asn Ala Glu Asp Val	1165	1170	1175	
gtg tgc tgc tca atg tct tac tct tgg aca ggc gca ctc gtc acc ccg				16323
Val Cys Cys Ser Met Ser Tyr Ser Trp Thr Gly Ala Leu Val Thr Pro	1180	1185	1190	
tgc gcc gcg gaa gaa cag aaa ctg ccc atc aat gca cta agc aac tcg				16371
Cys Ala Ala Glu Glu Gln Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser	1195	1200	1205	
ttg cta cgt cac cac aat ttg gtg tat tcc acc acc tca cgc agt gct				16419
Leu Leu Arg His His Asn Leu Val Tyr Ser Thr Ser Arg Ser Ala				

1210	1215	1220	1225	
tgc caa agg cag aag aaa gtc aca ttt gac aga ctg caa gtt ctg gac Cys Gln Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp				16467
1230	1235	1240		
agc cat tac cag gac gta ctc aag gag gtt aaa gca gcg gcg tca aaa Ser His Tyr Gln Asp Val Leu Lys Glu Val Lys Ala Ala Ser Lys				16515
1245	1250	1255		
gtg aag gct aac ttg cta tcc gta gag gaa gct tgc agc ctg acg ccc Val Lys Ala Asn Leu Leu Ser Val Glu Glu Ala Cys Ser Leu Thr Pro				16563
1260	1265	1270		
cca cac tca gcc aaa tcc aag ttt ggt tat ggg gca aaa gac gtc cgt Pro His Ser Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg				16611
1275	1280	1285		
tgc cat gcc aga aag gcc gta acc cac atc aac tcc gtg tgg aaa gac Cys His Ala Arg Lys Ala Val Thr His Ile Asn Ser Val Trp Lys Asp				16659
1290	1295	1300	1305	
ctt ctg gaa gac aat gta aca cca ata gac act acc atc atg gct aag Leu Leu Glu Asp Asn Val Thr Pro Ile Asp Thr Thr Ile Met Ala Lys				16707
1310	1315	1320		
aac gag gtt ttc tgc gtt cag cct gag aag ggg ggt cgt aag cca gct Asn Glu Val Phe Cys Val Gln Pro Glu Lys Gly Arg Lys Pro Ala				16755
1325	1330	1335		
cgt ctc atc gtg ttc ccc gat ctg ggc gtg cgc gtg tgc gaa aag atg Arg Leu Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met				16803
1340	1345	1350		
gct ttg tac gac gtg gtt aca aag ctc ccc ttg gcc gtg atg gga agc Ala Leu Tyr Asp Val Val Thr Lys Leu Pro Leu Ala Val Met Gly Ser				16851
1355	1360	1365		
tcc tac gga ttc caa tac tca cca gga cag cgg gtt gaa ttc ctc gtg Ser Tyr Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val				16899
1370	1375	1380	1385	
caa gcg tgg aag tcc aag aaa acc cca atg ggg ttc tcg tat gat acc Gln Ala Trp Lys Ser Lys Thr Pro Met Gly Phe Ser Tyr Asp Thr				16947
1390	1395	1400		
cgc tgc ttt gac tcc aca gtc act gag agc gac atc cgt acg gag gag Arg Cys Phe Asp Ser Thr Val Thr Glu Ser Asp Ile Arg Thr Glu Glu				16995
1405	1410	1415		
gca atc tac caa tgt tgt gac ctc gac ccc caa gcc cgc gtg gcc atc Ala Ile Tyr Gln Cys Cys Asp Leu Asp Pro Gln Ala Arg Val Ala Ile				17043
1420	1425	1430		
aag tcc ctc acc gag agg ctt tat gtt ggg ggc cct ctt acc aat tca Lys Ser Leu Thr Glu Arg Leu Tyr Val Gly Gly Pro Leu Thr Asn Ser				17091
1435	1440	1445		

/ \

agg ggg gag aac tgc ggc tat cgc agg tgc cgc gcg agc ggc gta ctg Arg Gly Glu Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu	1450 1455 1460 1465	17139
aca act agc tgt ggt aac acc ctc act tgc tac atc aag gcc cg <sup>g</sup> gca Thr Thr Ser Cys Gly Asn Thr Leu Thr Cys Tyr Ile Lys Ala Arg Ala	1470 1475 1480	17187
gcc tgt cga gcc gca ggg ctc cag gac tgc acc atg ctc gtg tgt ggc Ala Cys Arg Ala Ala Gly Leu Gln Asp Cys Thr Met Leu Val Cys Gly	1485 1490 1495	17235
gac gac tta gtc gtt atc tgt gaa agc gcg ggg gtc cag gag gac gcg Asp Asp Leu Val Val Ile Cys Glu Ser Ala Gly Val Gln Glu Asp Ala	1500 1505 1510	17283
gcg agc ctg aga gcc ttc acg gag gct atg acc agg tac tcc gcc ccc Ala Ser Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro	1515 1520 1525	17331
cct ggg gac ccc cca caa cca gaa tac gac ttg gag ctc ata aca tca Pro Gly Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser	1530 1535 1540 1545	17379
tgc tcc tcc aac gtg tca gtc gcc cac gac ggc gct gga aag agg gtc Cys Ser Ser Asn Val Ser Val Ala His Asp Gly Ala Gly Lys Arg Val	1550 1555 1560	17427
tac tac ctc acc cgt gac cct aca acc ccc ctc gcg aga gct gcg tgg Tyr Tyr Leu Thr Arg Asp Pro Thr Pro Leu Ala Arg Ala Ala Trp	1565 1570 1575	17475
gag aca gca aga cac act cca gtc aat tcc tgg cta ggc aac ata atc Glu Thr Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile	1580 1585 1590	17523
atg ttt gcc ccc aca ctg tgg gcg agg atg ata ctg atg acc cat ttc Met Phe Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe	1595 1600 1605	17571
ttt agc gtc ctt ata gcc agg gac cag ctt gaa cag gcc ctc gat tgc Phe Ser Val Leu Ile Ala Arg Asp Gln Leu Glu Gln Ala Leu Asp Cys	1610 1615 1620 1625	17619
gag atc tac ggg gcc tgc tac tcc ata gaa cca ctg gat cta cct cca Glu Ile Tyr Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro Pro	1630 1635 1640	17667
atc att caa aga ctc cat ggc ctc agc gca ttt tca ctc cac agt tac Ile Ile Gln Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr	1645 1650 1655	17715
tct cca ggt gaa atc aat agg gtg gcc gca tgc ctc aga aaa ctt ggg Ser Pro Gly Glu Ile Asn Arg Val Ala Ala Cys Leu Arg Lys Leu Gly	1660 1665 1670	17763
gta ccg ccc ttg cga gct tgg aga cac cgg gcc cg <sup>g</sup> agc gtc cg <sup>c</sup> gct		17811

1675                    1680                    1685

agg ctt ctg gcc aga gga ggc agg gct gcc ata tgt ggc aag tac ctc      17859  
 Arg Leu Leu Ala Arg Gly Gly Arg Ala Ala Ile Cys Gly Lys Tyr Leu  
 1690                    1695                    1700                    1705

ttc aac tgg gca gta aga aca aag ctc aaa ctc act cca ata gcg gcc      17907  
 Phe Asn Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Ala Ala  
 1710                    1715                    1720

gct ggc cag ctg gac ttg tcc ggc tgg ttc acg gct ggc tac agc ggg      17955  
 Ala Gly Gln Leu Asp Leu Ser Gly Trp Phe Thr Ala Gly Tyr Ser Gly  
 1725                    1730                    1735

gga gac att tat cac agc gtg tct cat gcc cg<sup>g</sup> ccc cgc tgg atc tgg      18003  
 Gly Asp Ile Tyr His Ser Val Ser His Ala Arg Pro Arg Trp Ile Trp  
 1740                    1745                    1750

ttt tgc cta ctc ctg ctt gct gca ggg gta ggc atc tac ctc ctc ccc      18051  
 Phe Cys Leu Leu Leu Ala Ala Gly Val Gly Ile Tyr Leu Leu Pro  
 1755                    1760                    1765

aac cga tgaaggttgg ggttaaacact ccggcctaaa aaaaaaaaaaa aatctagaac      18107  
 Asn Arg  
 1770

ccgagtcgac tttgtccca ctgtactttt agtcgtaca aaatacaata tactttcat 18167  
 ttctccgtaa acaacatgtt ttcccatgtt atatcctttt ctatTTTcg ttccgttacc 18227  
 aactttcacac atactttata tagctattca ttctataca ctaaaaaact aagacaattt 18287  
 taatTTTgct gcctgccata tttcaatttg ttataaattc ctataattta tcctattagt 18347  
 agctaaaaaa agatgaatgt gaatcgaatc ctaagagaat tggatctgat ccacaggacg 18407  
 ggtgtggtcg ccatgatcgc gtagtcgata gtggctcaa gtgcgaacg gagcaggact 18467  
 gggcggcggc caaagcggtc ggacagtgtt ccgagaacgg gtgcgcataa aaattgcata 18527  
 aacgcataata ggcgttagcag cacggcataa tgactggcga tgctgtcgga atggacgata 18587  
 tcccccaaga ggccggcag taccggcata accaaggcta tgcctacagc atccagggtg 18647  
 acggtgccga ggatgacgat gagcgcattt ttagattca tacacggtgc ctgactgcgt 18707  
 tagcaattta actgtgataa actaccgcat taaagctttt tctttccaat tttttttttt 18767  
 tcgtcattat aaaaatcatt acgaccgaga ttccggta ataactgata taatTTAattt 18827  
 gaagctctaa tttgtgagtt tagtatacat gcatttactt ataatacagt ttttttagttt 18887  
 tgctggccgc atcttctcaa atatgcTTCC cagcctgctt ttctgtaaacg ttcaccctct 18947  
 accttagcat cccttccctt tgcaaataagt cctttccaa caataataat gtcagatcct 19007

002247-000  
gtagagacca catcatccac gtttctatac tggtagccca atgcgtctcc cttgtcatct 19067  
aaacccacac cgggtgtcat aatcaaccaa tcgtaaccc tcattcttcc acccatgtct 19127  
cttgagcaa taaagccat aacaaaatct ttgtcgctt tcgcaatgtc aacagtaccc 19187  
ttatgtatatt ctccagtaga tagggagccc ttgcatgaca attctgctaa catcaaaagg 19247  
cctctaggtt cctttttac ttcttctgcc gcctgcttca aaccgctaac aatacctggg 19307  
cccaccacac cgtgtgcatt cgtaatgtct gccattctg ctattctgta tacacccgca 19367  
gagtaactgca atttgactgt attaccaatg tcagcaattt ttctgtcttc gaagagtaaa 19427  
aaattgtact tggcggataa tgcccttagc ggcttaactg tgccctccat ggaaaaatca 19487  
gtcaagatat ccacatgtgt ttttagtaaa caaatttgg gacctaattgc ttcaactaac 19547  
tccagtaatt ctttgtgg acgaacatcc aatgaagcac acaagtttgt ttgctttcg 19607  
tgcatgatataatagctt ggcagcaaca ggactaggat gagtagcagc acgttcctta 19667  
tatgttagctt tcgacatgtat tatcttcgt ttccgtcagg ttttgttct gtgcagttgg 19727  
gttaagaata ctggcaatt tcatgtttct tcaacactac atatgcgtat atataccaat 19787  
ctaagtctgt gtccttcct tcgttctcc ttctgttcgg agattaccga atcaaaaaaaa 19847  
tttcaaggaa accgaaatca aaaaaaagaa taaaaaaaaaa atgatgaatt gaaaagctta 19907  
tcgat 19912

<210> 9  
<211> 1771  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: pd.deltaNS3NS5

<400> 9  
Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro  
1 5 10 15

Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His  
20 25 30

Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly  
35 40 45

Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly  
50 55 60

Cys Ser Gly Gly Ala Tyr Asp Ile Ile Cys Asp Glu Cys His Ser  
65 70 75 80

Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Ala  
                   85                         90                         95  
  
 Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala Thr Ala Thr Pro Pro  
                   100                     105                         110  
  
 Gly Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser  
                   115                     120                         125  
  
 Thr Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val  
                   130                     135                         140  
  
 Ile Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Lys Lys Cys  
                   145                     150                         155                 160  
  
 Asp Glu Leu Ala Ala Lys Leu Val Ala Leu Gly Ile Asn Ala Val Ala  
                   165                     170                         175  
  
 Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val  
                   180                     185                         190  
  
 Val Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe  
                   195                     200                         205  
  
 Asp Ser Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe  
                   210                     215                         220  
  
 Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr Ile Thr Leu Pro Gln Asp  
                   225                     230                         235                 240  
  
 Ala Val Ser Arg Thr Gln Arg Arg Gly Arg Thr Gly Arg Gly Lys Pro  
                   245                     250                         255  
  
 Gly Ile Tyr Arg Phe Val Ala Pro Gly Glu Arg Pro Ser Gly Met Phe  
                   260                     265                         270  
  
 Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp Ala Gly Cys Ala Trp Tyr  
                   275                     280                         285  
  
 Glu Leu Thr Pro Ala Glu Thr Thr Val Arg Leu Arg Ala Tyr Met Asn  
                   290                     295                         300  
  
 Thr Pro Gly Leu Pro Val Cys Gln Asp His Leu Glu Phe Trp Glu Gly  
                   305                     310                         315                 320  
  
 Val Phe Thr Gly Leu Thr His Ile Asp Ala His Phe Leu Ser Gln Thr  
                   325                     330                         335  
  
 Lys Gln Ser Gly Glu Asn Leu Pro Tyr Leu Val Ala Tyr Gln Ala Thr  
                   340                     345                         350  
  
 Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser Trp Asp Gln Met Trp  
                   355                     360                         365  
  
 Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His Gly Pro Thr Pro Leu  
                   370                     375                         380

Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Ile Thr Leu Thr His Pro  
385 390 395 400

Val Thr Lys Tyr Ile Met Thr Cys Met Ser Ala Asp Leu Glu Val Val  
405 410 415

Thr Ser Thr Trp Val Leu Val Gly Gly Val Leu Ala Ala Leu Ala Ala  
420 425 430

Tyr Cys Leu Ser Thr Gly Cys Val Val Ile Val Gly Arg Val Val Leu  
435 440 445

Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu Val Leu Tyr Arg Glu  
450 455 460

Phe Asp Glu Met Glu Glu Cys Ser Gln His Leu Pro Tyr Ile Glu Gln  
465 470 475 480

Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys Ala Leu Gly Leu Leu  
485 490 495

Gln Thr Ala Ser Arg Gln Ala Glu Val Ile Ala Pro Ala Val Gln Thr  
500 505 510

Asn Trp Gln Lys Leu Glu Thr Phe Trp Ala Lys His Met Trp Asn Phe  
515 520 525

Ile Ser Gly Ile Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn  
530 535 540

Pro Ala Ile Ala Ser Leu Met Ala Phe Thr Ala Ala Val Thr Ser Pro  
545 550 555 560

Leu Thr Thr Ser Gln Thr Leu Leu Phe Asn Ile Leu Gly Gly Trp Val  
565 570 575

Ala Ala Gln Leu Ala Ala Pro Gly Ala Ala Thr Ala Phe Val Gly Ala  
580 585 590

Gly Leu Ala Gly Ala Ala Ile Gly Ser Val Gly Leu Gly Lys Val Leu  
595 600 605

Ile Asp Ile Leu Ala Gly Tyr Gly Ala Gly Val Ala Gly Ala Leu Val  
610 615 620

Ala Phe Lys Ile Met Ser Gly Glu Val Pro Ser Thr Glu Asp Leu Val  
625 630 635 640

Asn Leu Leu Pro Ala Ile Leu Ser Pro Gly Ala Leu Val Val Gly Val  
645 650 655

Val Cys Ala Ala Ile Leu Arg Arg His Val Gly Pro Gly Glu Gly Ala  
660 665 670

Val Gln Trp Met Asn Arg Leu Ile Ala Phe Ala Ser Arg Gly Asn His  
675 680 685

Val Ser Pro Thr His Tyr Val Pro Glu Ser Asp Ala Ala Ala Arg Val  
690 695 700

Thr Ala Ile Leu Ser Ser Leu Thr Val Thr Gln Leu Leu Arg Arg Leu  
705 710 715 720

His Gln Trp Ile Ser Ser Glu Cys Thr Thr Pro Cys Ser Gly Ser Trp  
725 730 735

Leu Arg Asp Ile Trp Asp Trp Ile Cys Glu Val Leu Ser Asp Phe Lys  
740 745 750

Thr Trp Leu Lys Ala Lys Leu Met Pro Gln Leu Pro Gly Ile Pro Phe  
755 760 765

Val Ser Cys Gln Arg Gly Tyr Lys Gly Val Trp Arg Gly Asp Gly Ile  
770 775 780

Met His Thr Arg Cys His Cys Gly Ala Glu Ile Thr Gly His Val Lys  
785 790 795 800

Asn Gly Thr Met Arg Ile Val Gly Pro Arg Thr Cys Arg Asn Met Trp  
805 810 815

Ser Gly Thr Phe Pro Ile Asn Ala Tyr Thr Thr Gly Pro Cys Thr Pro  
820 825 830

Leu Pro Ala Pro Asn Tyr Thr Phe Ala Leu Trp Arg Val Ser Ala Glu  
835 840 845

Glu Tyr Val Glu Ile Arg Gln Val Gly Asp Phe His Tyr Val Thr Gly  
850 855 860

Met Thr Thr Asp Asn Leu Lys Cys Pro Cys Gln Val Pro Ser Pro Glu  
865 870 875 880

Phe Phe Thr Glu Leu Asp Gly Val Arg Leu His Arg Phe Ala Pro Pro  
885 890 895

Cys Lys Pro Leu Leu Arg Glu Glu Val Ser Phe Arg Val Gly Leu His  
900 905 910

Glu Tyr Pro Val Gly Ser Gln Leu Pro Cys Glu Pro Glu Pro Asp Val  
915 920 925

Ala Val Leu Thr Ser Met Leu Thr Asp Pro Ser His Ile Thr Ala Glu  
930 935 940

Ala Ala Gly Arg Arg Leu Ala Arg Gly Ser Pro Pro Ser Val Ala Ser  
945 950 955 960

Ser Ser Ala Ser Gln Leu Ser Ala Pro Ser Leu Lys Ala Thr Cys Thr  
965 970 975

Ala Asn His Asp Ser Pro Asp Ala Glu Leu Ile Glu Ala Asn Leu Leu  
980 985 990

Trp Arg Gln Glu Met Gly Gly Asn Ile Thr Arg Val Glu Ser Glu Asn  
995 1000 1005

Lys Val Val Ile Leu Asp Ser Phe Asp Pro Leu Val Ala Glu Glu Asp  
1010 1015 1020

Glu Arg Glu Ile Ser Val Pro Ala Glu Ile Leu Arg Lys Ser Arg Arg  
025 1030 1035 1040

Phe Ala Gln Ala Leu Pro Val Trp Ala Arg Pro Asp Tyr Asn Pro Pro  
1045 1050 1055

Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr Glu Pro Pro Val Val His  
1060 1065 1070

Gly Cys Pro Leu Pro Pro Lys Ser Pro Pro Val Pro Pro Pro Arg  
1075 1080 1085

Lys Lys Arg Thr Val Val Leu Thr Glu Ser Thr Leu Ser Thr Ala Leu  
1090 1095 1100

Ala Glu Leu Ala Thr Arg Ser Phe Gly Ser Ser Ser Thr Ser Gly Ile  
105 1110 1115 1120

Thr Gly Asp Asn Thr Thr Ser Ser Glu Pro Ala Pro Ser Gly Cys  
1125 1130 1135

Pro Pro Asp Ser Asp Ala Glu Ser Tyr Ser Ser Met Pro Pro Leu Glu  
1140 1145 1150

Gly Glu Pro Gly Asp Pro Asp Leu Ser Asp Gly Ser Trp Ser Thr Val  
1155 1160 1165

Ser Ser Glu Ala Asn Ala Glu Asp Val Val Cys Cys Ser Met Ser Tyr  
1170 1175 1180

Ser Trp Thr Gly Ala Leu Val Thr Pro Cys Ala Ala Glu Glu Gln Lys  
185 1190 1195 1200

Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu Arg His His Asn Leu  
1205 1210 1215

Val Tyr Ser Thr Thr Ser Arg Ser Ala Cys Gln Arg Gln Lys Lys Val  
1220 1225 1230

Thr Phe Asp Arg Leu Gln Val Leu Asp Ser His Tyr Gln Asp Val Leu  
1235 1240 1245

Lys Glu Val Lys Ala Ala Ser Lys Val Lys Ala Asn Leu Leu Ser  
1250 1255 1260

Val Glu Glu Ala Cys Ser Leu Thr Pro Pro His Ser Ala Lys Ser Lys  
265 1270 1275 1280

Phe Gly Tyr Gly Ala Lys Asp Val Arg Cys His Ala Arg Lys Ala Val  
1285 1290 1295

Thr His Ile Asn Ser Val Trp Lys Asp Leu Leu Glu Asp Asn Val Thr  
1300 1305 1310

Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu Val Phe Cys Val Gln  
1315 1320 1325

Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu Ile Val Phe Pro Asp  
1330 1335 1340

Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu Tyr Asp Val Val Thr  
345 1350 1355 1360

Lys Leu Pro Leu Ala Val Met Gly Ser Ser Tyr Gly Phe Gln Tyr Ser  
1365 1370 1375

Pro Gly Gln Arg Val Glu Phe Leu Val Gln Ala Trp Lys Ser Lys Lys  
1380 1385 1390

Thr Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe Asp Ser Thr Val  
1395 1400 1405

Thr Glu Ser Asp Ile Arg Thr Glu Glu Ala Ile Tyr Gln Cys Cys Asp  
1410 1415 1420

Leu Asp Pro Gln Ala Arg Val Ala Ile Lys Ser Leu Thr Glu Arg Leu  
425 1430 1435 1440

Tyr Val Gly Gly Pro Leu Thr Asn Ser Arg Gly Glu Asn Cys Gly Tyr  
1445 1450 1455

Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr Ser Cys Gly Asn Thr  
1460 1465 1470

Leu Thr Cys Tyr Ile Lys Ala Arg Ala Ala Cys Arg Ala Ala Gly Leu  
1475 1480 1485

Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp Leu Val Val Ile Cys  
1490 1495 1500

Glu Ser Ala Gly Val Gln Glu Asp Ala Ala Ser Leu Arg Ala Phe Thr  
505 1510 1515 1520

Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp Pro Pro Gln Pro  
1525 1530 1535

Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser Asn Val Ser Val  
1540 1545 1550

Ala His Asp Gly Ala Gly Lys Arg Val Tyr Tyr Leu Thr Arg Asp Pro  
1555 1560 1565

Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr Ala Arg His Thr Pro  
1570 1575 1580

Val Asn Ser Trp Leu Gly Asn Ile Ile Met Phe Ala Pro Thr Leu Trp  
585 1590 1595 1600

Ala Arg Met Ile Leu Met Thr His Phe Phe Ser Val Leu Ile Ala Arg  
1605 1610 1615

Asp Gln Leu Glu Gln Ala Leu Asp Cys Glu Ile Tyr Gly Ala Cys Tyr  
1620 1625 1630

Ser Ile Glu Pro Leu Asp Leu Pro Pro Ile Ile Gln Arg Leu His Gly  
1635 1640 1645

Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro Gly Glu Ile Asn Arg  
1650 1655 1660

Val Ala Ala Cys Leu Arg Lys Leu Gly Val Pro Pro Leu Arg Ala Trp  
665 1670 1675 1680

Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu Leu Ala Arg Gly Gly  
1685 1690 1695

Arg Ala Ala Ile Cys Gly Lys Tyr Leu Phe Asn Trp Ala Val Arg Thr  
1700 1705 1710

Lys Leu Lys Leu Thr Pro Ile Ala Ala Ala Gly Gln Leu Asp Leu Ser  
1715 1720 1725

Gly Trp Phe Thr Ala Gly Tyr Ser Gly Gly Asp Ile Tyr His Ser Val  
1730 1735 1740

Ser His Ala Arg Pro Arg Trp Ile Trp Phe Cys Leu Leu Leu Ala  
745 1750 1755 1760

Ala Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg  
1765 1770

<210> 10

<211> 19798

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:  
pd.deltaNS3NS5.pj

<220>

<221> CDS

<222> (12679) .. (17991)

<400> 10

atcgatccata ccccttgcgc taaagaagta tatgtgccta ctaacgcttg tctttgtctc 60

tgtcaactaaa cactggatta ttactccag atacttattt tggactaatt taaatgattt 120

cggatcaacg ttcttaatat cgctgaatct tccacaattt atgaaaggtag ctaggaagag 180

gaattggat aaagttttt ttttgtaaa tctcgaagta tactcaaacg aatttagtat 240

tttctcagtg atctcccaga tgctttcacc ctcacttaga agtgctttaa gcatttttt 300

actgtggcta tttcccttat ctgcttcttc cgatgattcg aactgttaatt gcaaactact 360  
tacaatatca gtgatatacg attgatgttt ttgtccatag taaggaaataa ttgtaaattc 420  
ccaagcagga atcaatttct ttaatgaggc ttccagaatt gttgctttt gcgtcttgta 480  
tttaaactgg agtgatttat tgacaatatac gaaactcagc gaattgctta tgatagtatt 540  
atagctcatg aatgtggctc tcttgattgc tgccccgtta tgtgtaatca tccaacataa 600  
ataggtagt tcagcagcac ataatgctat tttctcacct gaaggtctt caaaccttcc 660  
cacaactga cgaacaagca ccttaggtgg tgtttacat aatataatcaa attgtggcat 720  
gcttagcgcc gatcttgtgt gcaattgata tctagttca actactctat ttatcttgta 780  
tcttcagta ttcaaacacg ctaactcgaa aaactaactt taattgtcct gtttgtctcg 840  
cgttcttcg aaaaatgcac cggccgcgca ttatttgtac tgcggaaaata attggtaactg 900  
cggtatcttc atttcatatt taaaaatgc acctttgctg ctttccctta atttttagac 960  
ggcccgagg ttctgttgc ggtactatct tgtataaaaa agttgtttt acatgtgatc 1020  
tgcacagatt ttataatgta ataagcaaga atacattatac aaacgaacaa tactggtaaa 1080  
agaaaaccaa aatggacgac attgaaacag ccaagaatct gacggtaaaa gcacgtacag 1140  
cttatacgct ctggatgta tgccgttgtt ttattgaaat gattgctcct gatgtagata 1200  
ttgatataga gagtaaacgt aagtctgatg agctactctt tccaggatata gtcataaggc 1260  
ccatggaatc ttcacaacc ggtaggccgt atggcttga ttctagcgca gaagattcca 1320  
gcttatcttc tgactccagt gctgaggtaa tttgcctgc tgcgaagatg gttaggaaa 1380  
ggtttgattc gattgaaat ggtatgctct ttcacaaga agcaagtcag gctgccatag 1440  
atttgcgt acagaataac aagctgttag acaatagaaa gcaactatac aaatctattg 1500  
ctataataat aggaagattg cccgagaaag acaagaagag agctaccgaa atgctcatga 1560  
gaaaaatgga ttgtacacag ttattagtc caccagctcc aacggaagaa gatgttatga 1620  
agctcgtaag cgtcgtaacc caattgctta cttagttcc accagatcgta caagctgctt 1680  
taataggtga tttattcatc ccggaatctc taaaggatata attcaatagt ttcaatgaac 1740  
tggcggcaga gaatcgtaa cagcaaaaaa agagtgagtt ggaaggaagg actgaagtga 1800  
accatgctaa tacaaatgaa gaagttccct ccaggcgaac aagaagtaga gacacaaatg 1860  
caagaggagc atataaatta caaaacacca tcactgaggg ccctaaagcg gttcccacga 1920  
aaaaaaggag agtagcaacg agggtaaggg gcagaaaatc acgtaataact tctagggtat 1980  
gatccaatat caaaggaaat gatagcatg aaggatgaga ctaatccaaat tgaggagtgg 2040

DNA sequence

cagcatatag aacagctaaa gggtagtgct gaaggaagca tacgataccc cgcatggaat 2100  
gggataatat cacaggaggt actagactac ctttcattc acataaaatag acgcataataa 2160  
gtacgcattt aagcataaac acgcactatg ccgttcttct catgtatata tatatacagg 2220  
caacacgcag atataggtgc gacgtgaaca gtgagctgta tgcgcgcagc tcgcgttgca 2280  
tttcggaaag cgctcgaaaa cggaaacgct ttgaagttcc tattccgaag ttcctattct 2340  
ctagaaagta taggaacttc agagcgctt tgaaaaccaa aagcgctctg aagacgcact 2400  
ttcaaaaaac caaaaacgca ccggactgta acgagctact aaaatattgc gaataccgct 2460  
tccacaaaca ttgctaaaaa gtatctctt gctatatac tctgtgctat atccctatat 2520  
aacctaccca tccacccccc gtccttggaa cttgcatacta aactcgacct ctacatcaac 2580  
aggcttccaa tgctttcaaa attttactgt caagtagacc catacggttg taatatgctg 2640  
ctcttcataaa tgtaagctta tctttatcga atcgtgtgaa aaactactac cgcgataaac 2700  
ctttacggtt ccctgagatt gaatttagttc cttagtata tgataacaaga cacttttgaa 2760  
ctttgtacga cgaattttga gtttcggccat cctctggcta tttccaattt tcctgtcgcc 2820  
tattatctcc gcctcagttt gatcttccgc ttcaactgc cattttcac ataatgaatc 2880  
tatttcaccc cacaatcctt catccgcctc cgtatcttgt tccgttaaac tattgacttc 2940  
atgttgtaca ttgttttagtt cacgagaagg gtcctttca ggcggtagct cctgatctcc 3000  
tatatgacct ttatcctgtt ctctttccac aaacttagaa atgtattcat gaattatgga 3060  
gcacctaata acattttca aggccggagaa gtttggcca gatgccaaat atgcttgaca 3120  
tgaaaacgtg agaatgaatt tagtattatt gtgatattct gaggcaattt tattataatc 3180  
tcgaagataa gagaagaatg cagtgcaccc ttttgcata aatggagatt ccatgtatct 3240  
aaaaaaatcg cctttaggcc ttctgatacc cttttccctg cggttttagcg tgcctttac 3300  
attaatatct aaaccctctc cgatggtgcc ctttaactga ctaataaaatg caaccgatata 3360  
aaactgtgat aattctgggt gatttatgat tcgatcgaca attgtattgt acactagtgc 3420  
aggatcagggc caatccagtt cttttcaat taccgggttg tcgtctgtat tcagtgatcg 3480  
tccaacaaat gcaaatgcta acgttttgta tttcttataa ttgtcaggaa ctggaaaaagt 3540  
cccccttgcgtc gtctcgatataa cacaatctact ttcatcgatcc accatagggtt ggaagtgcgt 3600  
cataatacat tgcttaatac aagcaagcag tctctcgcca ttcatatcc agttatcc 3660  
cattacagct gatgtcattt gatatacaggc ctgtaaaaat ctatctgtta cagaaggttt 3720  
tcgcggtttt tataaacaataa actttcgta cgaaatcgag caatcacccc agctgcgtat 3780

ttggaaattc gggaaaaagt agagcaacgc gagttgcatt ttttacacca taatgcata 3840  
ttaacttcga gaagggatta aggctaattt cactagtatg tttcaaaaac ctcaatctgt 3900  
ccattgaatg ccttataaaa cagctataga ttgcataaga gagtagtca ctcaatgctt 3960  
tttgtcaaag cttactgtatg atgatgtgtc tacttcagg cggtctgtt gtaaggagaa 4020  
tgacattata aagctggcac tttagaattcc acggactata gactatacta gtatactccg 4080  
tctactgtac gatacacttc cgctcaggc cttgccttt aacgaggcct taccactctt 4140  
ttgttactct attgatccag ctcagcaaag gcagtgtat ctaagattct atttcgca 4200  
tgtagtaaaa ctagctagac cgagaaagag actagaaatg caaaaggcac ttctacaatg 4260  
gctgccatca ttattatccg atgtgacgct gcatttttt tttttttttt tttttttttt 4320  
ttttttttt tttttttttt tttttggta caaatatcat aaaaaaagag aatctttta 4380  
agcaaggatt ttcttaactt ctccggcgcac agcatcacccg acttcggtgg tactgttgaa 4440  
accacctaaa tcaccagttc tgataacctgc atccaaaacc ttttaactg catcttcaat 4500  
ggcttaccc tcttcaggca agttcaatga caattcaac atcattgcag cagacaagat 4560  
agtggcgata gggttgacct tattcttgg caaatctgga gcggaaaccat ggcattggc 4620  
gtacaaacca aatgcggtgt tcttgctgg caaagaggcc aaggacgcag atggcaacaa 4680  
acccaaggag cctggataa cggaggcttc atcggagatg atatcaccaa acatgttgct 4740  
ggtgattata ataccattt agtgggttgg gttcttaact aggatcatgg cggcagaatc 4800  
aatcaattga tggtaactt tcaatgttagt gaattcggtt ttgatggttt cctccacagt 4860  
tttctccat aatcttgaag aggccaaaac attagctta tccaggacc aaataggcaa 4920  
tggggctca tggtaactt ccattttttt ggcattttt gtgattttt gcacttctgg 4980  
aacgggttat tggtaactt cccaggcgcac accatcacca tcgttttctt ttcttttacc 5040  
aaagtaaata cctccacta attctcttac aacaacgaag tcagttaccc tagaaattt 5100  
tggcttgatt ggagataagt ctaaaagaga gtcggatgca aagttacatg gtcttaagtt 5160  
ggcgtacaat tggtaactt tacggatttt tagtaaacct tggtaactt taacactacc 5220  
ggtaaaaaat ttaggaccac ccacaggcacc taacaaaacg gcatcaggcct tcttggaggc 5280  
ttccagcgcc tcatctggaa gtggaaacacc tggatgtatc atggcggcacc caccaattaa 5340  
atgattttcg aaatcgaact tggatgtatc acggacatca gaaatagtt taagaacctt 5400  
aatggcttcg gctgtgattt ctggatccac gtggtcaccc ggcaaaacga cgatcttctt 5460  
agggcagac attacaatgg tttttttttt aaatatatat aaaaaaaaaa aaaaaaaaaa 5520

aaaaaaaaaa atgcagcttc tcaatgatat tcgaatacgc ttggaggaga tacagcctaa 5580  
tatccgacaa actgtttac agatttacga tcgtacttgt tacccatcat tgaattttga 5640  
acatccgaac ctggagttt tccctgaaac agatagtata ttgtaccctg tataataata 5700  
tatagtctag cgcttacgg aagacaatgt atgtattcg ttccctggag aaactattgc 5760  
atctattgca taggtaatct tgcacgtcgc atccccggtt cattttctgc gtttccatct 5820  
tgcaactcaa tagcatatct ttgttaacga agcatctgtg cttcattttg tagaacaaaa 5880  
atgcaacgcg agagcgctaa ttttcaaac aaagaatctg agctgcattt ttacagaaca 5940  
gaaatgcaac gcgaaagcgc tattttacca acgaagaatc tgtgcttcat ttttgtaaaa 6000  
caaaaatgca acgcgagagc gctaattttt caaacaaaga atctgagctg catttttaca 6060  
gaacagaaat gcaacgcgag agcgctattt taccaacaaa gaatctatac ttctttttg 6120  
ttctacaaaaa atgcaccccg agagcgctat ttttctaaca aagcatctta gattactttt 6180  
tttctcctt gtgcgtctta taatgcagtc tcttgataac ttttgcact gtaggtccgt 6240  
taaggtaga agaaggctac tttgggtgtct attttctctt ccataaaaaa agcctgactc 6300  
cacttcccgc gtttactgat tactagcgaa gctgggggtg catttttca agataaaggc 6360  
atccccgatt atattctata ccgatgtgga ttgcgcatac tttgtgaaca gaaagtgata 6420  
gcgttgatga ttcttcattt gtcagaaaaat tatgaacggt ttcttctatt ttgtctctat 6480  
atactacgta taggaatgt ttacatttc gtattgtttt cgattcactc tatgaatagt 6540  
tcttactaca attttttgt ctaaagagta atactagaga taaacataaa aaatgttagag 6600  
gtcgagttt gatgcaagtt caaggagcga aagggtggatg ggtaggttat atagggatat 6660  
agcacagaga tatatacgaa agagatactt ttgagcaatg tttgtggaag cggatttcgc 6720  
aatatttttag tagctcgta cagtcgggtg cgtttttgtt ttttgaaag tgcgtcttca 6780  
gagcgctttt ggtttcaaa agcgctctga agttcctata ctttctagag aataggaact 6840  
tcggaatagg aacttcaaag cgtttccgaa aacgagcgct tccgaaaatg caacgcgagc 6900  
tgcgcacata cagctcactg ttcacgtcgc acctatatct gcgtgttgc ttttatata 6960  
tatacatgag aagaacggca tagtgcgtgt ttatgcttaa atgcgtactt atatgcgtct 7020  
atttatgttag gatgaaaggt agtctagtagtac ctcctgtat attatccat tccatgcggg 7080  
gtatcgatgtat cttcccttcag cactaccctt tagtgcgtct atatgcgtcc actcctcaat 7140  
tggatttagtc tcacccctca atgctatcat ttcccttgcatttggatcat atgcatacgta 7200  
ccgagaaaact agtgcgaagt agtgcgtcagg tattgcgtttt atctgatgag tatacggtt 7260

00000000000000000000000000000000

cctggccacg gcagaagcac gcttatcgct ccaattccc acaacat tag 7320  
taggccttc attgaaagaa atgaggtcat caaatgtctt ccaatgtgag attttggcc 7380  
atttttata gcaaagattg aataaggcgc attttctt aaagctttat tgtacgatct 7440  
gactaagtta tctttataa attggatttc ctgttattt cttgaagaat tgccggtcct 7500  
atttactcg ttttagactg gttcagaatt cctcaaaaat tcatccaaat atacaagtgg 7560  
atcgatgata agctgtcaaa catgagaatt cttgaagacg aaagggcctc gtgatacgcc 7620  
tattttata ggttaatgtc atgataataa tggttctt gacgtcaggt ggcactttc 7680  
ggggaaatgt gcgcggaacc cctatttgtt tattttcta aatacattca aatatgtatc 7740  
cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg aagagtgatga 7800  
gtattcaaca tttccgtgtc gcccttattt ccttttgc ggcattttgc cttcctgttt 7860  
ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag 7920  
tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagttt cgccccgaag 7980  
aacgtttcc aatgatgagc acttttaaag ttctgctatg tggcgcggta ttatccgtg 8040  
ttgacgcccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttgggtt 8100  
agtactcacc agtcacagaa aagcatctt cggatggcat gacagtaaga gaattatgca 8160  
gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca acgatcgag 8220  
gaccgaagga gctaaccgct ttttgcaca acatggggta tcatgtact cgccttgatc 8280  
gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg 8340  
cagcaatggc aacaacgttg cgcaaaactat taactggcgactacttact ctagctccc 8400  
ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg 8460  
cccttccggc tggctggttt attgctgata aatctggagc cggtagcgt gggtctcg 8520  
gtatcattgc agcaactgggg ccagatggta agccctcccg tatcgttagtt atctacacga 8580  
cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac 8640  
tgattaagca ttggtaactg tcagaccaag tttactcata tatacttttag attgatttaa 8700  
aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgcataat ctcatgacca 8760  
aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag 8820  
gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaaca aaaaaaccac 8880  
cgctaccagc ggtgggttgt ttgcggatc aagagctacc aactctttt ccgaaggtaa 8940  
ctggcttcag cagagcgcag ataccaaata ctgtccttct agtgtagccg tagtaggccc 9000

accacttcaa gaactctgta gcaccgccta catacctcg tctgctaata ctgttaccag 9060  
tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac 9120  
cgatataaggc gcagcggctcg ggctgaacgg ggggttcgtg cacacagccc agcttggagc 9180  
gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc 9240  
ccgaaggggag aaaggcggac aggtatccgg taagcggcag ggtcggaaaca ggagagcgca 9300  
cgagggagct tccaggggaa aacgccttgtt atctttatag tcctgtcggg tttcgccacc 9360  
tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg 9420  
ccagcaacgc ggcctttta cggttcctgg cctttgtg gcctttgct cacatgttct 9480  
ttcctgcgtt atcccctgat tctgtggata accgtattac cgccttgag tgagctgata 9540  
ccgctcgccg cagccgaacg accgagcgcga gcgagtcagt gagcgaggaa gcggaagagc 9600  
gcctgatgct gtattttctc cttacgcattc tgtgcgttat ttcacaccgc atatggtgca 9660  
ctctcagtac aatctgctct gatgccgcat agttaagcca gtatacactc cgctatcgct 9720  
acgtgactgg gtcatggctg cgccccgaca cccgccaaca cccgctgacg cgccctgacg 9780  
ggcttgcgtc ctcccccgtat ccgcttacag acaagctgtg accgtctccg ggagctgcat 9840  
gtgtcagagg ttttccacgt catcaccgaa acgcgcgagg cagctgcggt aaagctcatc 9900  
acgctggctg tgaagcgatt cacagatgtc tgcgttca tccgcgttcca gtcgttgag 9960  
tttctccaga agcgttaatg tctggcttct gataaagcgg gccatgttaa gggcggttt 10020  
ttcctgtttt gtcactgatg cctccgtgtt aaaaaaaaaaaaaaaaatgatgatgat 10080  
accgatgaaa cgagagagga tgctcacgt acgggtact gatgtgaac atgcccgggtt 10140  
actggAACGT tgcgtgggtt aacaactggc ggtatggatg cggcgggacc agagaaaaat 10200  
cactcagggt caatgccagc gttcgtaa tacagatgtt ggtgttccac aggtagcca 10260  
gcagcatcct gcgtgcaga tccggaaacat aatggtgcag ggcgtact tccgcgtttc 10320  
cagactttac gaaacacgga aaccgaagac cattcatgtt gttgctcagg tcgcagacgt 10380  
tttgcagcag cagtcgttcc acgttcgttc gcgtatcggt gattcattct gctaaccagt 10440  
aaggcaaccc cgccagccta gcccgggtctt caacgacagg agcacgtca tgccacccg 10500  
tggccaggac ccaacgctgc ccgagatcg ccgcgtgcgg ctgctggaga tggcggacgc 10560  
gatggatatg ttctgcaag gttgggttttgcgttccaca gtttccgca agaattgatt 10620  
ggctccaatt cttggagtgg tgaatccgtt agcgaggtgc cgccggcttc cattcaggc 10680  
gaggtggccc ggctccatgc accgcgcacgc aacgcggggaa ggcagacaag gtatagggcg 10740

gcccctacaa tccatccaa cccgttccat gtgctcgccg aggcggcata aatcgccgtg 10800  
acgatcagcg gtccaatgat cgaagttagg ctggtaagag ccgcgagcga tccttgaagc 10860  
tgtccctgat ggtcgcatc tacctgcctg gacagcatgg cctgcaacgc gggcatcccg 10920  
atgccgcccgg aagcgagaag aatcataatg gggaaaggcca tccagcctcg cgtcgcgaac 10980  
gccagcaaga cgtagccccag cgctcgccgg gccatcccg cgataatggc ctgcttctcg 11040  
ccgaaacgtt tggtgccggg accagtgacg aaggotttag cgagggcgtg caagattccg 11100  
aataccgcaa gcgacaggcc gatcatcg tcgcgtccagc gaaagcggtc ctcgcccggaaa 11160  
atgaccgcaga gcgctgcccgg cacctgtcct acgagttgca tgataaagaa gacagtcata 11220  
agtgcggcga cgatagtcat gccccgcgcc cacccggagg agctgactgg gttgaaggct 11280  
ctcaaggcga tcggtcgagg atccttcaat atgcgcacat acgctgttat gttcaaggc 11340  
ccttcgttta agaacgaaag cggtcttcct ttgagggat gttcaagtt gttcaaatct 11400  
atcaaatttg caaatccccca gtctgtatct agagcggtga atcggtgatg cgatttggta 11460  
attaaattga tggtgtcacc attaccaggc ttagatatac caatggcaaa ctgagcacaa 11520  
caataccagt ccggatcaac tggcaccatc tctcccgtag tctcatctaa tttttcttcc 11580  
ggatgaggtt ccagatatac cgcaacaccc ttattatggt ttccctgagg gaataataga 11640  
atgtcccatt cgaaatcacc aattctaaac ctggcgaat tgtatccgg gtttggtaac 11700  
tcgttccagt caggaatgtt ccacgtgaag ctatcttcca gcaaagtctc cacttcttca 11760  
tcaaattgtg gagaatactc ccaatgtct tatctatggg acttccggg aacacagtc 11820  
cgataacttcc caattcgctc tcagagctca ttgtttgtt gaagagacta atcaaagaat 11880  
cgaaaaattaa tatcttaact gatagttga tcaaaggccg aaaacgtagg 11940  
ggcaaacaataa cggaaaaatc gttccaaa tttctgtatc ccaagaactc taaccagtct 12000  
tatctaaaaaa ttgccttatg atccgtctc cccgttacag cctgtgtac tgattaatcc 12060  
tgccttctca atcaccattc taatgtttta attaaggat tttgtcttca ttaacggctt 12120  
tcgctcataa aaatgttatg acgtttgcc cgccggcggg aaaccatcca cttcacgaga 12180  
ctgatctcct ctgcccggac accggggcatc tccaaacttat aagttggaga aataagagaa 12240  
tttcagattt agagaatgaa aaaaaaaaaac ctttagttca taggtccatt ctcttagcgc 12300  
aactacagag aacagggggca caaacaggca aaaaacgggc acaacctcaa tggagtgtatc 12360  
caacctgcct ggagtaaatg atgacacacaag gcaattgacc cacgcatgta tctatctcat 12420  
tttcttacac cttcttattac cttctgctc ctctgattt gaaaaagctg aaaaaaaaaagg 12480

ttgaaaccag ttccctgaaa ttattcccct acttgactaa taatgtata aagacggtag 12540  
 gtattgattt gtttctgtt aatctatttc tttaaacttct taaattctac ttttatagtt 12600  
 agtctttttt ttatgtttaa aacaccaaga acttagttc gaataaacac acataaaacaa 12660  
 acaagcttac aaaacaaa atg gct gca tat gca gct cag ggc tat aag gtg 12711  
                   Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val  
                   1              5              10  
 cta gta ctc aac ccc tct gtt gct gca aca ctg ggc ttt ggt gct tac 12759  
 Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr  
                   15              20              25  
 atg tcc aag gct cat ggg atc gat cct aac atc agg acc ggg gtg aga 12807  
 Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg  
                   30              35              40  
 aca att acc act ggc agc ccc atc acg taa tcc acc tac ggc aag ttc 12855  
 Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe  
                   45              50              55  
 ctt gcc gac ggc ggg tgc tcg ggg ggc gct tat gac ata ata att tgt 12903  
 Leu Ala Asp Gly Gly Cys Ser Gly Ala Tyr Asp Ile Ile Ile Cys  
                   60              65              70              75  
 gac gag tgc cac tcc acg gat gcc aca tcc atc ttg ggc att ggc act 12951  
 Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr  
                   80              85              90  
 gtc ctt gac caa gca gag act gcg ggg ggc aga ctg gtt gtg ctc gcc 12999  
 Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala  
                   95              100              105  
 acc gcc acc cct ccg ggc tcc gtc act gtg ccc cat ccc aac atc gag 13047  
 Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn Ile Glu  
                   110              115              120  
 gag gtt gct ctg tcc acc acc gga gag atc cct ttt tac ggc aag gct 13095  
 Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala  
                   125              130              135  
 atc ccc ctc gaa gta atc aag ggg ggg aga cat ctc atc ttc tgt cat 13143  
 Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu Ile Phe Cys His  
                   140              145              150              155  
 tca aag aag aag tgc gac gaa ctc gcc gca aag ctg gtc gca ttg ggc 13191  
 Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Ala Leu Gly  
                   160              165              170  
 atc aat gcc gtg gcc tac tac cgc ggt ctt gac gtg tcc gtc atc ccg 13239  
 Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile Pro  
                   175              180              185  
 acc agc ggc gat gtt gtc gtc gtc gca acc gat gcc ctc atg acc ggc 13287  
 Thr Ser Gly Asp Val Val Val Ala Thr Asp Ala Leu Met Thr Gly  
                   190              195              200

1

tat acc ggc gac ttc gac tcg gtg ata gac tgc aat acg tgt gtc acc		13335	
Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys Val Thr			
205	210	215	
cag aca gtc gat ttc agc ctt gac cct acc ttc acc att gag aca atc		13383	
Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr Ile			
220	225	230	235
acg ctc ccc caa gat gct gtc tcc cgc act caa cgt cgg ggc agg act		13431	
Thr Leu Pro Gln Asp Ala Val Ser Arg Thr Gln Arg Arg Gly Arg Thr			
240	245	250	
ggc agg ggg aag cca ggc atc tac aga ttt tgt gca ccg ggg gag cgc		13479	
Gly Arg Gly Lys Pro Gly Ile Tyr Arg Phe Val Ala Pro Gly Glu Arg			
255	260	265	
ccc tcc ggc atg ttc gac tcg tcc gtc ctc tgt gag tgc tat gac gca		13527	
Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp Ala			
270	275	280	
ggc tgt gct tgg tat gag ctc acg ccc gag act aca gtt agg cta		13575	
Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Thr Val Arg Leu			
285	290	295	
cga gcg tac atg aac acc ccg ggg ctt ccc tgt tgc cag gac cat ctt		13623	
Arg Ala Tyr Met Asn Thr Pro Gly Leu Pro Val Cys Gln Asp His Leu			
300	305	310	315
gaa ttt tgg gag ggc gtc ttt aca ggc ctc act cat ata gat gcc cac		13671	
Glu Phe Trp Glu Gly Val Phe Thr Gly Leu Thr His Ile Asp Ala His			
320	325	330	
ttt cta tcc cag aca aag cag agt ggg gag aac ctt cct tac ctg gta		13719	
Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Leu Pro Tyr Leu Val			
335	340	345	
gcg tac caa gcc acc gtg tgc gct agg gct caa gcc cct ccc cca tcg		13767	
Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser			
350	355	360	
tgg gac cag atg tgg aag tgt ttg att cgc ctc aag ccc acc ctc cat		13815	
Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His			
365	370	375	
ggg cca aca ccc ctg cta tac aga ctg ggc gct gtt cag aat gaa atc		13863	
Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Ile			
380	385	390	395
acc ctg acg cac cca gtc acc aaa tac atc atg aca tgc atg tcg gcc		13911	
Thr Leu Thr His Pro Val Thr Lys Tyr Ile Met Thr Cys Met Ser Ala			
400	405	410	
gac ctg gag gtc acg agc acc tgg gtg ctc gtt ggc ggc gtc ctg		13959	
Asp Leu Glu Val Val Thr Ser Thr Trp Val Leu Val Gly Gly Val Leu			
415	420	425	
gct gct ttg gcc gcg tat tgc ctg tca aca ggc tgc gtg gtc ata gtg		14007	

DRAFT GENOME SEQUENCING

DRAFT

Ala Ala Leu Ala Ala Tyr Cys Leu Ser Thr Gly Cys Val Val Ile Val			
430	435	440	
ggc agg gtc gtc ttg tcc ggg aag ccg gca atc ata cct gac agg gaa			14055
Gly Arg Val Val Leu Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu			
445	450	455	
gtc ctc tac cga gag ttc gat gag atg gaa gag tgc tct cag cac tta			14103
Val Leu Tyr Arg Glu Phe Asp Glu Met Glu Glu Cys Ser Gln His Leu			
460	465	470	475
ccg tac atc gag caa ggg atg atg ctc gco gag cag ttc aag cag aag			14151
Pro Tyr Ile Glu Gln Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys			
480	485	490	
gcc ctc ggc ctc ctg cag acc gcg tcc cct cag gca gag gtt atc gcc			14199
Ala Leu Gly Leu Leu Gln Thr Ala Ser Arg Gln Ala Glu Val Ile Ala			
495	500	505	
cct gct gtc cag acc aac tgg caa aaa ctc gag acc ttc tgg gcg aag			14247
Pro Ala Val Gln Thr Asn Trp Gln Lys Leu Glu Thr Phe Trp Ala Lys			
510	515	520	
cat atg tgg aac ttc atc agt ggg ata caa tac ttg gcg ggc ttg tca			14295
His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala Gly Leu Ser			
525	530	535	
acg ctg cct ggt aac ccc gcc att gct tca ttg atg gct ttt aca gct			14343
Thr Leu Pro Gly Asn Pro Ala Ile Ala Ser Leu Met Ala Phe Thr Ala			
540	545	550	555
gct gtc acc agc cca cta acc act agc caa acc ctc ctc ttc aac ata			14391
Ala Val Thr Ser Pro Leu Thr Thr Ser Gln Thr Leu Leu Phe Asn Ile			
560	565	570	
ttg ggg ggg tgg gtg gct gcc cag ctc gcc gcc ccc ggt gcc gct act			14439
Leu Gly Gly Trp Val Ala Ala Gln Leu Ala Ala Pro Gly Ala Ala Thr			
575	580	585	
gcc ttt gtg ggc gct ggc tta gct ggc gcc gcc atc ggc agt gtt gga			14487
Ala Phe Val Gly Ala Gly Leu Ala Gly Ala Ala Ile Gly Ser Val Gly			
590	595	600	
ctg ggg aag gtc ctc ata gac atc ctt gca ggg tat ggc gcg ggc gtg			14535
Leu Gly Lys Val Leu Ile Asp Ile Leu Ala Gly Tyr Gly Ala Gly Val			
605	610	615	
gcg gga gct ctt gtg gca ttc aag atc atg agc ggt gag gtc ccc tcc			14583
Ala Gly Ala Leu Val Ala Phe Lys Ile Met Ser Gly Glu Val Pro Ser			
620	625	630	635
acg gag gac ctg gtc aat cta ctg ccc gcc atc ctc tcg ccc gga gcc			14631
Thr Glu Asp Leu Val Asn Leu Leu Pro Ala Ile Leu Ser Pro Gly Ala			
640	645	650	
ctc gta gtc ggc gtg gtc tgc gca gca ata ctg cgc cgg cac gtt ggc			14679
Leu Val Val Gly Val Val Cys Ala Ala Ile Leu Arg Arg His Val Gly			

655

660

665

ccg ggc gag ggg gca gtg cag tgg atg aac cgg	ctg ata gcc ttc gcc	14727
Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg	Leu Ile Ala Phe Ala	
670	675	680
tcc cgg ggg aac cat gtt tcc ccc acg cac tac	gtg ccg gag agc gat	14775
Ser Arg Gly Asn His Val Ser Pro Thr His Tyr	Val Pro Glu Ser Asp	
685	690	695
gca gct gcc cgc gtc act gcc ata ctc agc	agc ctc act gta acc cag	14823
Ala Ala Ala Arg Val Thr Ala Ile Leu Ser	Ser Leu Thr Val Thr Gln	
700	705	710
715		
ctc ctg agg cga ctg cac cag tgg ata agc	tcg gag tgt acc act cca	14871
Leu Leu Arg Arg Leu His Gln Trp Ile Ser	Ser Glu Cys Thr Thr Pro	
720	725	730
tgc tcc ggt tcc tgg cta agg gac atc	tgg gac tgg ata tgc gag gtg	14919
Cys Ser Gly Ser Trp Leu Arg Asp	Ile Trp Asp Trp Ile Cys Glu Val	
735	740	745
ttg agc gac ttt aag acc tgg cta aaa	gct aag ctc atg cca cag ctg	14967
Leu Ser Asp Phe Lys Thr Trp	Leu Lys Ala Lys Leu Met Pro Gln Leu	
750	755	760
cct ggg atc ccc ttt gtg tcc tgc cag cgc	ggg tat aag ggg gtc tgg	15015
Pro Gly Ile Pro Phe Val Ser Cys Gln Arg	Gly Tyr Lys Gly Val Trp	
765	770	775
cga ggg gac ggc atc atg cac act	cgc tgc cac tgt gga gct gag atc	15063
Arg Gly Asp Gly Ile Met His Thr Arg Cys	His Cys Gly Ala Glu Ile	
780	785	790
795		
act gga cat gtc aaa aac ggg acq atg agg	atc gtc ggt cct agg acc	15111
Thr Gly His Val Lys Asn Gly Thr Met Arg	Ile Val Gly Pro Arg Thr	
800	805	810
tgc agg aac atg tgg agt ggg ac	ttc ccc att aat gcc tac acc acg	15159
Cys Arg Asn Met Trp Ser Gly Thr	Phe Pro Ile Asn Ala Tyr Thr	
815	820	825
ggc ccc tgt acc ccc ctt cct	gcg ccg aac tac acg ttc gcg cta tgg	15207
Gly Pro Cys Thr Pro Leu Pro	Ala Pro Asn Tyr Thr Phe Ala Leu Trp	
830	835	840
agg gtg tct gca gag gaa	tao gtg gag ata agg cag gtg ggg gac ttc	15255
Arg Val Ser Ala Glu Glu	Tyr Val Glu Ile Arg Gln Val Gly Asp Phe	
845	850	855
cac tac gtg acg ggt atg act	act gac aat ctt aaa tgc ccg tgc cag	15303
His Tyr Val Thr Gly Met Thr	Asp Asn Leu Lys Cys Pro Cys Gln	
860	865	870
875		
gtc cca tcg ccc gaa ttt	ttc aca gaa ttg gac ggg gtg cgc cta cat	15351
Val Pro Ser Pro Glu Phe	Phe Thr Glu Leu Asp Gly Val Arg Leu His	
880	885	890

DRAFT 6/24/2000

agg ttt gcg ccc ccc tgc aag ccc ttg ctg cg <del>g</del> gag gag gta tca ttc		15399
Arg Phe Ala Pro Pro Cys Lys Pro Leu Leu Arg Glu Glu Val Ser Phe		
895 900 905		
aga gta gga ctc cac gaa tac ccg gta ggg tc <del>g</del> caa tta cct tgc gag		15447
Arg Val Gly Leu His Glu Tyr Pro Val Gly Ser Gln Leu Pro Cys Glu		
910 915 920		
ccc gaa ccg gac gtg gcc gtg ttg acg tcc at <del>g</del> ctc act gat ccc tcc		15495
Pro Glu Pro Asp Val Ala Val Leu Thr Ser Met Leu Thr Asp Pro Ser		
925 930 935		
cat ata aca gca gag g <del>c</del> ggg cga agg ttg g <del>c</del> agg gga tca ccc		15543
His Ile Thr Ala Glu Ala Ala Gly Arg Arg Leu Ala Arg Gly Ser Pro		
940 945 950 955		
ccc tct gtg gcc agc tcc tc <del>g</del> gct agc cag cta tcc gct cca tct ctc		15591
Pro Ser Val Ala Ser Ser Ala Ser Gln Leu Ser Ala Pro Ser Leu		
960 965 970		
aag gca act tgc acc gct aac cat gac tcc cct gat gct gag ctc ata		15639
Lys Ala Thr Cys Thr Ala Asn His Asp Ser Pro Asp Ala Glu Leu Ile		
975 980 985		
gag gcc aac ctc cta tgg agg cag gag at <del>g</del> ggc ggc aac atc acc agg		15687
Glu Ala Asn Leu Leu Trp Arg Gln Glu Met Gly Gly Asn Ile Thr Arg		
990 995 1000		
gtt gag tca gaa aac aaa gtg gtg att ctg gac tcc ttc gat ccg ctt		15735
Val Glu Ser Glu Asn Lys Val Val Ile Leu Asp Ser Phe Asp Pro Leu		
1005 1010 1015		
gtg g <del>c</del> gag gag gac gag cgg gag atc tcc gta ccc gca gaa atc ctg		15783
Val Ala Glu Glu Asp Glu Arg Glu Ile Ser Val Pro Ala Glu Ile Leu		
1020 1025 1030 1035		
cg <del>g</del> aag tct cgg aga ttc gcc cag gcc ctg ccc gtt tgg g <del>c</del> cg <del>g</del> ccg		15831
Arg Lys Ser Arg Arg Phe Ala Gln Ala Leu Pro Val Trp Ala Arg Pro		
1040 1045 1050		
gac tat aac ccc ccg cta gtg gag acg tgg aaa aag ccc gac tac gaa		15879
Asp Tyr Asn Pro Pro Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr Glu		
1055 1060 1065		
cca cct gtg gtc cat gg <del>c</del> tgc cc <del>g</del> ctt cca cct cca aag tcc cct cct		15927
Pro Pro Val Val His Gly Cys Pro Leu Pro Pro Pro Lys Ser Pro Pro		
1070 1075 1080		
gtg cct cc <del>g</del> cct cc <del>g</del> aag aag cg <del>g</del> acg gtg gtc ctc act gaa tca acc		15975
Val Pro Pro Pro Arg Lys Lys Arg Thr Val Val Leu Thr Glu Ser Thr		
1085 1090 1095		
cta tct act gcc ttg gcc qag ctc gcc acc aga agc ttt ggc agc tcc		16023
Leu Ser Thr Ala Leu Ala Glu Leu Ala Thr Arg Ser Phe Gly Ser Ser		
1100 1105 1110 1115		
tca act tcc ggc att acg ggc gac aat acg aca aca tcc tct gag ccc		16071

80

Ser Thr Ser Gly Ile Thr Gly Asp Asn Thr Thr Thr Ser Ser Glu Pro			
1120	1125	1130	
gcc cct tct ggc tgc ccc ccc gac tcc gac got gag tcc tat tcc tcc			16119
Ala Pro Ser Gly Cys Pro Pro Asp Ser Asp Ala Glu Ser Tyr Ser Ser			
1135	1140	1145	
atg ccc ccc ctg gag ggg gag cct ggg gat ccg gat ctt agc gac ggg			16167
Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu Ser Asp Gly			
1150	1155	1160	
tca tgg tca acg gtc agt agt gag gcc aac gcg gag gat gtc gtg tgc			16215
Ser Trp Ser Thr Val Ser Ser Glu Ala Asn Ala Glu Asp Val Val Cys			
1165	1170	1175	
tgc tca atg tct tac tct tgg aca ggc gca ctc gtc acc ccg tgc gcc			16263
Cys Ser Met Ser Tyr Ser Trp Thr Gly Ala Leu Val Thr Pro Cys Ala			
1180	1185	1190	1195
gcg gaa gaa cag aaa ctg ccc atc aat gca cta agc aac tcg ttg cta			16311
Ala Glu Glu Gln Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu			
1200	1205	1210	
cgt cac cac aat ttg gtg tat tcc acc acc tca cgc agt gct tgc caa			16359
Arg His His Asn Leu Val Tyr Ser Thr Thr Ser Arg Ser Ala Cys Gln			
1215	1220	1225	
agg cag aag aaa gtc aca ttt gac aga ctg caa gtt ctg gac agc cat			16407
Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Ser His			
1230	1235	1240	
tac cag gac gta ctc aag gag gtt aaa gca gcg gcg tca aaa gtg aag			16455
Tyr Gln Asp Val Leu Lys Glu Val Lys Ala Ala Ser Lys Val Lys			
1245	1250	1255	
gct aac ttg cta tcc gta gag gaa gct tgc agc ctg acg ccc cca cac			16503
Ala Asn Leu Leu Ser Val Glu Glu Ala Cys Ser Leu Thr Pro Pro His			
1260	1265	1270	1275
tca gcc aaa tcc aag ttt ggt tat ggg gca aaa gac gtc cgt tgc cat			16551
Ser Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Cys His			
1280	1285	1290	
gcc aga aag gcc gta acc cac atc aac tcc gtg tgg aaa gac ctt ctg			16599
Ala Arg Lys Ala Val Thr His Ile Asn Ser Val Trp Lys Asp Leu Leu			
1295	1300	1305	
gaa gac aat gta aca cca ata gac act acc atc atg gct aag aac gag			16647
Glu Asp Asn Val Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu			
1310	1315	1320	
gtt ttc tgc gtt cag cct gag aag ggg ggt cgt aag cca gct cgt ctc			16695
Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu			
1325	1330	1335	
atc gtg ttc ccc gat ctg ggc gtg cgc gtg tgc gaa aag atg gct ttg			16743
Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu			

1340	1345	1350	1355	
tac gac gtg gtt aca aag ctc ccc ttg gcc gtg atg gga agc tcc tac Tyr Asp Val Val Thr Lys Leu Pro Leu Ala Val Met Gly Ser Ser Tyr				
1360	1365	1370		16791
gga ttc caa tac tca cca gga cag cgg gtt gaa ttc ctc gtg caa gcg Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Gln Ala				
1375	1380	1385		16839
tgg aag tcc aag aaa acc cca atg ggg ttc tcg tat gat acc cgc tgc Trp Lys Ser Lys Lys Thr Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys				
1390	1395	1400		16887
ttt gac tcc aca gtc act gag agc gac atc cgt acg gag gag gca atc Phe Asp Ser Thr Val Thr Glu Ser Asp Ile Arg Thr Glu Glu Ala Ile				
1405	1410	1415		16935
tac caa tgt tgt gac ctc gac ccc caa gcc cgc gtg gcc atc aag tcc Tyr Gln Cys Cys Asp Leu Asp Pro Gln Ala Arg Val Ala Ile Lys Ser				
1420	1425	1430	1435	16983
ctc acc gag agg ctt tat gtt ggg ggc cct ctt acc aat tca agg ggg Leu Thr Glu Arg Leu Tyr Val Gly Gly Pro Leu Thr Asn Ser Arg Gly				
1440	1445	1450		17031
gag aac tgc ggc tat cgc agg tgc cgc gcg agc ggc gta ctg aca act Glu Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr				
1455	1460	1465		17079
agc tgt ggt aac acc ctc act tgc tac atc aag gcc cgg gca gcc tgt Ser Cys Gly Asn Thr Leu Thr Cys Tyr Ile Lys Ala Arg Ala Ala Cys				
1470	1475	1480		17127
cga gcc gca ggg ctc cag gac tgc acc atg ctc gtg tgt ggc gac gac Arg Ala Ala Gly Leu Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp				
1485	1490	1495		17175
tta gtc gtt atc tgt gaa agc gcg ggg gtc cag gag gac gcg gcg agc Leu Val Val Ile Cys Glu Ser Ala Gly Val Gln Glu Asp Ala Ala Ser				
1500	1505	1510	1515	17223
ctg aga gcc ttc acg gag gct atg acc agg tac tcc gcc ccc cct ggg Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly				
1520	1525	1530		17271
gac ccc cca caa cca gaa tac gac ttg gag ctc ata aca tca tgc tcc Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser				
1535	1540	1545		17319
tcc aac gtg tca gtc gcc cac gac ggc gct gga aag agg gtc tac tac Ser Asn Val Ser Val Ala His Asp Gly Ala Gly Lys Arg Val Tyr Tyr				
1550	1555	1560		17367
ctc acc cgt gac cct aca acc ccc ctc gcg aga gct gcg tgg gag aca Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr				
1565	1570	1575		17415

gca aga cac act cca gtc aat tcc tgg cta ggc aac ata atc atg ttt 17463  
 Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Phe  
 1580 1585 1590 1595  
  
 gcc ccc aca ctg tgg gcg agg atg ata ctg atg acc cat ttc ttt agc 17511  
 Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser  
 1600 1605 1610  
  
 gtc ctt ata gcc agg gac cag ctt gaa cag gcc ctc gat tgc gag atc 17559  
 Val Leu Ile Ala Arg Asp Gln Leu Glu Gln Ala Leu Asp Cys Glu Ile  
 1615 1620 1625  
  
 tac ggg gcc tgc tac tcc ata gaa cca ctg gat cta cct cca atc att 17607  
 Tyr Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro Pro Ile Ile  
 1630 1635 1640  
  
 caa aga ctc cat ggc ctc agc gca ttt tca ctc cac agt tac tct cca 17655  
 Gln Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro  
 1645 1650 1655  
  
 ggt gaa atc aat agg gtg gcc gca tgc ctc aga aaa ctt ggg gta ccg 17703  
 Gly Glu Ile Asn Arg Val Ala Ala Cys Leu Arg Lys Leu Gly Val Pro  
 1660 1665 1670 1675  
  
 ccc ttg cga gct tgg aga cac cgg gcc cgg agc gtc cgc gct agg ctt 17751  
 Pro Leu Arg Ala Trp Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu  
 1680 1685 1690  
  
 ctg gcc aga gga ggc agg gct gcc ata tgt ggc aag tac ctc ttc aac 17799  
 Leu Ala Arg Gly Gly Arg Ala Ala Ile Cys Gly Lys Tyr Leu Phe Asn  
 1695 1700 1705  
  
 tgg gca gta aga aca aag ctc aaa ctc act cca ata gcg gcc gct ggc 17847  
 Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Ala Ala Gly  
 1710 1715 1720  
  
 cag ctg gac ttg tcc ggc tgg ttc acg gct ggc tac agc ggg gga gac 17895  
 Gln Leu Asp Leu Ser Gly Trp Phe Thr Ala Gly Tyr Ser Gly Gly Asp  
 1725 1730 1735  
  
 att tat cac agc gtg tct cat gcc cgg ccc cgc tgg atc tgg ttt tgc 17943  
 Ile Tyr His Ser Val Ser His Ala Arg Pro Arg Trp Ile Trp Phe Cys  
 1740 1745 1750 1755  
  
 cta ctc ctg ctt gct gca ggg gta ggc atc tac ctc ctc ccc aac cga 17991  
 Leu Leu Leu Ala Ala Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg  
 1760 1765 1770  
  
 tgaatagtcg actttgttcc cactgtactt ttagtcgtta caaaaatacaa tatactttc 18051  
  
 atttctccgt aaacaacatg ttttccatg taatatcctt ttctattttt cgttccgtta 18111  
  
 ccaactttac acatacttta tatagttatt cacttctata cactaaaaaa ctaagacaat 18171  
  
 tttaattttg ctgcctgcc a tatcaatt tgttataaat tcctataatt tatttttattt 18231  
  
 gtagctaaaa aaagatgaat gtgaatcgaa tcctaagaga attggatctg atccacagga 18291

000000000000000000000000

cgggtgtggt cgccatgatc gcgtagtcga tagtgctcc aagtagcgaa gcgagcagga 18351  
ctgggcggcg gccaaagcgg tcggacagtg ctccgagaac gggtgcgcat agaaaattgca 18411  
tcaacgcata tagcgctagc agcacgccc agtactggc gatgctgtcg gaatggacga 18471  
tatcccgcaa gaggcccggc agtaccggca taaccaagcc tatgcctaca gcatccaggg 18531  
tgacggtgcc gaggatgacg atgagcgcat tgtagattt catabacggt gcctgactgc 18591  
gttagcaatt taactgtgat aaactaccgc attaaagott tttcttcca atttttttt 18651  
tttcgtcatt ataaaaatca ttacgaccga gattccggg taataactga tataattaaa 18711  
ttgaagctct aatttgtgag tttagtatac atgcattac ttataataca gtttttagt 18771  
tttgctggcc gcatcttcctc aaatatgctt cccagcctgc tttctgtaa cgttcaccct 18831  
ctaccttagc atcccttcccc ttgcaaata gtccttcc aacaataata atgtcagatc 18891  
ctgttagagac cacatcatcc acgggtctat actgttgacc caatgcgtct cccttgtcat 18951  
ctaaacccac accgggtgtc ataatcaacc aatcgtaacc ttcatcttcc ccacccatgt 19011  
ctctttgagc aataaaagccg ataacaaaat ctgtcgct cttcgcaatg tcaacagtac 19071  
ccttagtata ttctccagta gatagggagc cttgtcatga caattctgct aacatcaaaa 19131  
ggcctctagg ttccttgtt acttcttcgt ccgcctgctt caaaccgcta acaatacctg 19191  
ggcccaccac accgtgtgca ttctgtatgt ctgcccattc tgctattctg tatacacccg 19251  
cagagtactg caatttgact gtattaccaa tgcagcaaa tttctgtct tcgaagagta 19311  
aaaaattgta ctggcggat aatgcctta gcggcttaac tgcgcctcc atggaaaaat 19371  
cagtcaagat atccacatgt gtttttagta aacaaatttt gggaccta atttcaacta 19431  
actccagtaa ttccttggtg gtacgaacat ccaatgaagc acacaagggtt gtttgccttt 19491  
cgtgcgtatgat attaaatagc ttggcagcaa caggactagg atgagtagca gcacgttcc 19551  
tatatgttagc ttctgacatg atttatcttgc gtttctgtca gttttttgtt ctgtgcagtt 19611  
gggttaagaa tactggcaa tttcatgttt cttcaacact acatatgcgt atatatacca 19671  
atctaagtct gtgccttc cttcggtt cttctgttc ggagattacc gaatcaaaaa 19731  
aatttcaagg aaaccgaaat caaaaaaaaaa aataaaaaaaaaa aaatgtgaa ttgaaaagct 19791  
tatcgat 19798

<210> 11  
<211> 1771  
<212> PRT  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:  
pd.deltaNS3NS5.pj

<400> 11

Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro  
1 5 10 15

Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His  
20 25 30

Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly  
35 40 45

Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly  
50 55 60

Cys Ser Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser  
65 70 75 80

Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Ala  
85 90 95

Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala Thr Ala Thr Pro Pro  
100 105 110

Gly Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser  
115 120 125

Thr Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val  
130 135 140

Ile Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Lys Cys  
145 150 155 160

Asp Glu Leu Ala Ala Lys Leu Val Ala Leu Gly Ile Asn Ala Val Ala  
165 170 175

Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val  
180 185 190

Val Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe  
195 200 205

Asp Ser Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe  
210 215 220

Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr Ile Thr Leu Pro Gln Asp  
225 230 235 240

Ala Val Ser Arg Thr Gln Arg Arg Gly Arg Thr Gly Arg Gly Lys Pro  
245 250 255

Gly Ile Tyr Arg Phe Val Ala Pro Gly Glu Arg Pro Ser Gly Met Phe  
260 265 270

Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp Ala Gly Cys Ala Trp Tyr

DRAFT 2/2000

275                    280                    285  
Glu Leu Thr Pro Ala Glu Thr Thr Val Arg Leu Arg Ala Tyr Met Asn  
290                    295                    300  
  
Thr Pro Gly Leu Pro Val Cys Gln Asp His Leu Glu Phe Trp Glu Gly  
305                    310                    315                    320  
  
Val Phe Thr Gly Leu Thr His Ile Asp Ala His Phe Leu Ser Gln Thr  
325                    330                    335  
  
Lys Gln Ser Gly Glu Asn Leu Pro Tyr Leu Val Ala Tyr Gln Ala Thr  
340                    345                    350  
  
Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser Trp Asp Gln Met Trp  
355                    360                    365  
  
Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His Gly Pro Thr Pro Leu  
370                    375                    380  
  
Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Ile Thr Leu Thr His Pro  
385                    390                    395                    400  
  
Val Thr Lys Tyr Ile Met Thr Cys Met Ser Ala Asp Leu Glu Val Val  
405                    410                    415  
  
Thr Ser Thr Trp Val Leu Val Gly Gly Val Leu Ala Ala Leu Ala Ala  
420                    425                    430  
  
Tyr Cys Leu Ser Thr Gly Cys Val Val Ile Val Gly Arg Val Val Leu  
435                    440                    445  
  
Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu Val Leu Tyr Arg Glu  
450                    455                    460  
  
Phe Asp Glu Met Glu Glu Cys Ser Gln His Leu Pro Tyr Ile Glu Gln  
465                    470                    475                    480  
  
Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys Ala Leu Gly Leu Leu  
485                    490                    495  
  
Gln Thr Ala Ser Arg Gln Ala Glu Val Ile Ala Pro Ala Val Gln Thr  
500                    505                    510  
  
Asn Trp Gln Lys Leu Glu Thr Phe Trp Ala Lys His Met Trp Asn Phe  
515                    520                    525  
  
Ile Ser Gly Ile Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn  
530                    535                    540  
  
Pro Ala Ile Ala Ser Leu Met Ala Phe Thr Ala Ala Val Thr Ser Pro  
545                    550                    555                    560  
  
Leu Thr Thr Ser Gln Thr Leu Leu Phe Asn Ile Leu Gly Gly Trp Val  
565                    570                    575  
  
Ala Ala Gln Leu Ala Ala Pro Gly Ala Ala Thr Ala Phe Val Gly Ala

580

585

590

Gly Leu Ala Gly Ala Ala Ile Gly Ser Val Gly Leu Gly Lys Val Leu  
 595 600 605

Ile Asp Ile Leu Ala Gly Tyr Gly Ala Gly Val Ala Gly Ala Leu Val  
 610 615 620

Ala Phe Lys Ile Met Ser Gly Glu Val Pro Ser Thr Glu Asp Leu Val  
 625 630 635 640

Asn Leu Leu Pro Ala Ile Leu Ser Pro Gly Ala Leu Val Val Gly Val  
 645 650 655

Val Cys Ala Ala Ile Leu Arg Arg His Val Gly Pro Gly Glu Gly Ala  
 660 665 670

Val Gln Trp Met Asn Arg Leu Ile Ala Phe Ala Ser Arg Gly Asn His  
 675 680 685

Val Ser Pro Thr His Tyr Val Pro Glu Ser Asp Ala Ala Ala Arg Val  
 690 695 700

Thr Ala Ile Leu Ser Ser Leu Thr Val Thr Gln Leu Leu Arg Arg Leu  
 705 710 715 720

His Gln Trp Ile Ser Ser Glu Cys Thr Thr Pro Cys Ser Gly Ser Trp  
 725 730 735

Leu Arg Asp Ile Trp Asp Trp Ile Cys Glu Val Leu Ser Asp Phe Lys  
 740 745 750

Thr Trp Leu Lys Ala Lys Leu Met Pro Gln Leu Pro Gly Ile Pro Phe  
 755 760 765

Val Ser Cys Gln Arg Gly Tyr Lys Gly Val Trp Arg Gly Asp Gly Ile  
 770 775 780

Met His Thr Arg Cys His Cys Gly Ala Glu Ile Thr Gly His Val Lys  
 785 790 795 800

Asn Gly Thr Met Arg Ile Val Gly Pro Arg Thr Cys Arg Asn Met Trp  
 805 810 815

Ser Gly Thr Phe Pro Ile Asn Ala Tyr Thr Thr Gly Pro Cys Thr Pro  
 820 825 830

Leu Pro Ala Pro Asn Tyr Thr Phe Ala Leu Trp Arg Val Ser Ala Glu  
 835 840 845

Glu Tyr Val Glu Ile Arg Gln Val Gly Asp Phe His Tyr Val Thr Gly  
 850 855 860

Met Thr Thr Asp Asn Leu Lys Cys Pro Cys Gln Val Pro Ser Pro Glu  
 865 870 875 880

Phe Phe Thr Glu Leu Asp Gly Val Arg Leu His Arg Phe Ala Pro Pro

DEPARTMENT OF

885                    890                    895

Cys Lys Pro Leu Leu Arg Glu Glu Val Ser Phe Arg Val Gly Leu His  
900                    905                    910

Glu Tyr Pro Val Gly Ser Gln Leu Pro Cys Glu Pro Glu Pro Asp Val  
915                    920                    925

Ala Val Leu Thr Ser Met Leu Thr Asp Pro Ser His Ile Thr Ala Glu  
930                    935                    940

Ala Ala Gly Arg Arg Leu Ala Arg Gly Ser Pro Pro Ser Val Ala Ser  
945                    950                    955                    960

Ser Ser Ala Ser Gln Leu Ser Ala Pro Ser Leu Lys Ala Thr Cys Thr  
965                    970                    975

Ala Asn His Asp Ser Pro Asp Ala Glu Leu Ile Glu Ala Asn Leu Leu  
980                    985                    990

Trp Arg Gln Glu Met Gly Gly Asn Ile Thr Arg Val Glu Ser Glu Asn  
995                    1000                    1005

Lys Val Val Ile Leu Asp Ser Phe Asp Pro Leu Val Ala Glu Glu Asp  
1010                    1015                    1020

Glu Arg Glu Ile Ser Val Pro Ala Glu Ile Leu Arg Lys Ser Arg Arg  
025                    1030                    1035                    1040

Phe Ala Gln Ala Leu Pro Val Trp Ala Arg Pro Asp Tyr Asn Pro Pro  
1045                    1050                    1055

Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr Glu Pro Pro Val Val His  
1060                    1065                    1070

Gly Cys Pro Leu Pro Pro Lys Ser Pro Pro Val Pro Pro Pro Arg  
1075                    1080                    1085

Lys Lys Arg Thr Val Val Leu Thr Glu Ser Thr Leu Ser Thr Ala Leu  
1090                    1095                    1100

Ala Glu Leu Ala Thr Arg Ser Phe Gly Ser Ser Ser Thr Ser Gly Ile  
105                    1110                    1115                    1120

Thr Gly Asp Asn Thr Thr Ser Ser Glu Pro Ala Pro Ser Gly Cys  
1125                    1130                    1135

Pro Pro Asp Ser Asp Ala Glu Ser Tyr Ser Ser Met Pro Pro Leu Glu  
1140                    1145                    1150

Gly Glu Pro Gly Asp Pro Asp Leu Ser Asp Gly Ser Trp Ser Thr Val  
1155                    1160                    1165

Ser Ser Glu Ala Asn Ala Glu Asp Val Val Cys Cys Ser Met Ser Tyr  
1170                    1175                    1180

Ser Trp Thr Gly Ala Leu Val Thr Pro Cys Ala Ala Glu Glu Gln Lys

185                    1190                    1195                    1200

Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu Arg His His Asn Leu  
                       1205                        1210                        1215

Val Tyr Ser Thr Thr Ser Arg Ser Ala Cys Gln Arg Gln Lys Lys Val  
                       1220                        1225                        1230

Thr Phe Asp Arg Leu Gln Val Leu Asp Ser His Tyr Gln Asp Val Leu  
                       1235                        1240                        1245

Lys Glu Val Lys Ala Ala Ser Lys Val Lys Ala Asn Leu Leu Ser  
                       1250                        1255                        1260

Val Glu Glu Ala Cys Ser Leu Thr Pro Pro His Ser Ala Lys Ser Lys  
                       265                        1270                        1275                        1280

Phe Gly Tyr Gly Ala Lys Asp Val Arg Cys His Ala Arg Lys Ala Val  
                       1285                        1290                        1295

Thr His Ile Asn Ser Val Trp Lys Asp Leu Leu Glu Asp Asn Val Thr  
                       1300                        1305                        1310

Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu Val Phe Cys Val Gln  
                       1315                        1320                        1325

Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu Ile Val Phe Pro Asp  
                       1330                        1335                        1340

Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu Tyr Asp Val Val Thr  
                       345                        1350                        1355                        1360

Lys Leu Pro Leu Ala Val Met Gly Ser Ser Tyr Gly Phe Gln Tyr Ser  
                       1365                        1370                        1375

Pro Gly Gln Arg Val Glu Phe Leu Val Gln Ala Trp Lys Ser Lys Lys  
                       1380                        1385                        1390

Thr Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe Asp Ser Thr Val  
                       1395                        1400                        1405

Thr Glu Ser Asp Ile Arg Thr Glu Glu Ala Ile Tyr Gln Cys Cys Asp  
                       1410                        1415                        1420

Leu Asp Pro Gln Ala Arg Val Ala Ile Lys Ser Leu Thr Glu Arg Leu  
                       425                        1430                        1435                        1440

Tyr Val Gly Gly Pro Leu Thr Asn Ser Arg Gly Glu Asn Cys Gly Tyr  
                       1445                        1450                        1455

Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr Ser Cys Gly Asn Thr  
                       1460                        1465                        1470

Leu Thr Cys Tyr Ile Lys Ala Arg Ala Ala Cys Arg Ala Ala Gly Leu  
                       1475                        1480                        1485

Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp Leu Val Val Ile Cys

1490	1495	1500
Glu Ser Ala Gly Val Gln Glu Asp Ala Ala		Ser Leu Arg Ala Phe Thr
505	1510	1515
Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro	Gly Asp Pro Pro Gln Pro	
1525	1530	1535
Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys	Ser Ser Asn Val Ser Val	
1540	1545	1550
Ala His Asp Gly Ala Gly Lys Arg Val Tyr	Tyr Leu Thr Arg Asp Pro	
1555	1560	1565
Thr Thr Pro Leu Ala Arg Ala Ala Trp	Glu Thr Ala Arg His Thr Pro	
1570	1575	1580
Val Asn Ser Trp Leu Gly Asn Ile Ile Met	Phe Ala Pro Thr Leu Trp	
585	1590	1595
Ala Arg Met Ile Leu Met Thr His Phe Phe	Ser Val Leu Ile Ala Arg	
1605	1610	1615
Asp Gln Leu Glu Gln Ala Leu Asp Cys	Glu Ile Tyr Gly Ala Cys Tyr	
1620	1625	1630
Ser Ile Glu Pro Leu Asp Leu Pro Pro	Ile Ile Gln Arg Leu His Gly	
1635	1640	1645
Leu Ser Ala Phe Ser Leu His Ser	Tyr Ser Pro Gly Glu Ile Asn Arg	
1650	1655	1660
Val Ala Ala Cys Leu Arg Lys Leu Gly	Val Pro Pro Leu Arg Ala Trp	
665	1670	1675
Arg His Arg Ala Arg Ser Val Arg Ala Arg	Leu Leu Ala Arg Gly Gly	
1685	1690	1695
Arg Ala Ala Ile Cys Gly Lys	Tyr Leu Phe Asn Trp Ala Val Arg Thr	
1700	1705	1710
Lys Leu Lys Leu Thr Pro Ile	Ala Ala Ala Gly Gln Leu Asp Leu Ser	
1715	1720	1725
Gly Trp Phe Thr Ala Gly Tyr	Ser Gly Gly Asp Ile Tyr His Ser Val	
1730	1735	1740
Ser His Ala Arg Pro Arg Trp	Ile Trp Phe Cys Leu Leu Leu Ala	
745	1750	1755
Ala Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg		1760
1765		1770

<210> 12  
<211> 20220  
<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:  
pd.delta.NS3NS5.pj.core121

<220>

<221> CDS

<222> (12679) .. (18354)

<400> 12

atcgatccta cccctgcgc taaagaagta tatgtgccta ctaacgcttgc 60  
tgtcaactaaa cactggatta ttactcccag atacattattt tggactaatt taaatgattt 120  
cggtcaacg ttcttaatat cgctgaatct tccacaattt atgaaaggtag ctaggaagag 180  
gaattggat aaagtttttgc tttttgtaaa tctcgaagta tactcaaacg aatttagtat 240  
tttctcagtg atctcccaga tgcttcacc ctcaacttaga agtgctttaa gcatttttt 300  
actgtggcta tttcccttat ctgcttcttc cgatgattcg aactgtaatt gcaaactact 360  
tacaatatacgtt gtgatatcag attgatgttt ttgtccatag taaggaataa ttgtaaattc 420  
ccaaggcagga atcaatttct ttaatgaggc ttccagaatt gttgctttt gcgcttgta 480  
tttaaactgg agtgatttat tgacaatatac gaaactcagc gaattgctta tgatagtatt 540  
atagctcatg aatgtggctc tcttgattgc tgccccgtta tgtgtaatca tccaaacataa 600  
ataggttagt tcagcagcac ataatgctat tttctcacct gaaggctttt caaaccttcc 660  
cacaactga cgaacaagca ccttaggtgg tgccccat aatatacataa attgtggcat 720  
gcttagcgcc gatcttgtgt gcaattgata tctagttca actactctat ttatcttgta 780  
tcttgagta ttcaaaacacg ctaactcgaa aaactaactt taattgtcct gtttgtctcg 840  
cggtcttcg aaaaatgcac cggccgcgc ttatttgatc tgcggaaata attggactg 900  
cggtatcttc atttcatatt taaaaatgc accttgctg ctccccctta atttttagac 960  
ggcccgagg ttcgtttgc ggtactatct tgtgataaaa agttgttttgc acatgtgatc 1020  
tgcacagatt ttataatgttataaataagcaaga atacattatc aaacgaacaa tactggtaaa 1080  
agaaaaaccaa aatggacgac attgaaacag ccaagaatct gacggtaaaa gcacgtacag 1140  
cttatacggttgc tgcggctgt ttattgaaat gattgctcct gatgtgatc 1200  
ttgatatagttataaataacgt aagtctgtatc agctactctt tccaggatattt gtcataaggc 1260  
ccatggaaatc tctcacaacc ggtaggccgt atggcttgc ttctagcgca gaagattcca 1320  
gctgtatcttc tgactccagt gctgaggtaa ttttgcctgc tgcgaagatc gttaaggaaa 1380

ggtttgcattt gattggaaat ggtatgctct cttcacaaga agcaagtcag gctgccatag 1440  
atttgatgctt acagaataac aagctgttag acaatagaaa gcaactatac aaatctattg 1500  
ctataataat aggaagattt cccgagaaaag acaagaagag agctaccgaa atgctcatga 1560  
gaaaaatgga ttgtacacag ttattagtcc caccagctcc aacggaagaa gatgttatga 1620  
agctcgtaag cgtcgttacc caattgctt cttagttcc accagatgtt caagctgctt 1680  
taataggta tttattcatc ccggaatctc taaaggatat attcaatagt ttcaatgaac 1740  
tggcggcaga gaatcgttt cagcaaaaaa agagttagttt ggaaggaagg actgaagtga 1800  
accatgctaa tacaaatgaa gaagttccctt ccaggcgaac aagaagttaga gacacaaatg 1860  
caagaggagc atataaatta caaaacacca tcactgaggg ccctaaagcg gttcccacga 1920  
aaaaaaggag agtagcaacg aggtaaggg gcagaaaaatc acgttaatact tctagggtat 1980  
gatccaatat caaaggaaat gatagcattt aaggatgaga ctaatccat tgaggagtgg 2040  
cagcatatac aacagctaaa ggtagtgctt gaaggaagaa tacgataacc cgcatttggat 2100  
gggataatat cacaggaggt actagactac ctccatctt acataaatac acgcattataa 2160  
gtacgcattt aagcataaac acgcactatg ccgttcttcatgtatata tatatacagg 2220  
caacacgcag atataaggta gacgtgaaca gtgagctgtt tgtgogcagc tcgcgttgc 2280  
tttcggaaag cgctcgaaaaa cggaaacgtt ttgaagttcc tattccgaag ttcctattct 2340  
ctagaaagta taggaacttc agagcgcttt tgaaaaccaa aagcgctctg aagacgcact 2400  
ttcaaaaaac caaaaacgca ccggactgtt acgactact aaaatattgc gaataccgct 2460  
tccacaaaca ttgctaaaaa gtatctttt gctatatac tctgtgtat atccctatata 2520  
aacctaccca tccacccatc gctccatgtt cttgcatttta aactcgacctt ctacatcaac 2580  
aggcttccaa tgctcttcaa attttactgtt caagtagacc catacggttg taatatgctg 2640  
ctcttcataa tgtaagctt tctttatcgtt atcgtgtgaa aaactactac cgcgataaac 2700  
ctttacgggtt ccctgagattt gaatttagttt cttagtata tgatataaga cactttgaa 2760  
ctttgtacgtt cgaattttgtt gtttcggccat cctctggctt tttccatata tcctgtcgcc 2820  
tattatctcc gcctcagttt gatcttccgtt ttcagactgc cattttcacataatgttac 2880  
tatttcaccc cacaatcctt catccgcctt cgcattttgtt tccgttaaac tattgacttc 2940  
atgttgtaca ttgttttagttt cacgagaagg gtcctttca ggcggtagctt cctgatctcc 3000  
tatatgaccc ttatcctgtt ctcttccac aaacttagaa atgtattcat gaattatgga 3060  
gcacctaata acattttca aggcggagaa gttttggcca gatgccaaat atgcttgaca 3120

tgaaaacgtg agaatgaatt tagtattatt gtgatattct gaggcaattt tattataatc 3180  
tcgaagataa gagaagaatg cagtgacctt tgtattgaca aatggagatt ccatgtatct 3240  
aaaaaaatacg cctttaggcc ttctgatacc ctccccctg cggttagcg tgcctttac 3300  
attaatatct aaaccctctc cgatggtggc cttaactga ctaataaatg caaccgat 3360  
aaactgttat aattctgggt gatttatgtat tcgatcgaca attgtattgt acactagtgc 3420  
aggatcagggc caatccagtt cttttcaat taccggtgtg tcgtctgtat tcagtagatg 3480  
tccaacaaat gcaa atgcta acgtttgtt tttcttataa ttgtcaggaa ctggaaaagt 3540  
cccccttgtc gtctcgatta cacacctact ttcatcgta accatagggtt ggaagtgctg 3600  
cataatacat tgcttaataac aagcaagcag tctctcgcc ttcatatccc agttatccc 3660  
cattacagct gatgtcattt tatatcagcg ctgtaaaaat ctatctgtt cagaaggttt 3720  
tcgcggtttt tataaacaactt actttcgta cgaaatcgag caatcacccc agctgcgtat 3780  
ttggaaattt gggaaaaagt agagcaacgc gagttgcatt ttttacacca taatgcattga 3840  
ttaacttcga gaagggatta aggctaattt cactagtatg tttcaaaaac ctcaatctgt 3900  
ccattgaatg ctttataaaaa cagctataga ttgcacagaa gagttagcta ctcaatgctt 3960  
tttgc当地 cttactgtatg atgatgtgtc tactttcagg cgggtctgtt gtaaggagaa 4020  
tgacattata aagctggcac tttagaattcc acggactata gactatacta gtatactccg 4080  
tctactgtac gataacttc cgctcaggc cttgtccctt aacgaggcct taccactctt 4140  
ttgttactctt attgatccag ctcagcaag gcagttgtat ctaagattct atcttcgcga 4200  
tgttagaaaa ctagctagac cgagaaagag actagaaatg caaaaggcac ttctacaatg 4260  
gctgccatca ttattatccg atgtgacgt gcattttttt tttttttttt tttttttttt 4320  
ttttttttt tttttttttt tttttggta caaatatcat aaaaaaaagag aatctttta 4380  
agcaaggatt ttcttaactt ctccggcgac agcatcaccg acttcgggtt tactgttgaa 4440  
accacctaata tcaccagttc tgataacctgc atccaaaacc ttttaactg catcttaat 4500  
ggctttacct tcttcaggca agttcaatga caatttcaac atcattgcag cagacaagat 4560  
agtggcgata gggttgacct tattctttgg caaatctgga gcggaaaccat ggcattggc 4620  
gtacaaaacca aatgcgggtt tcttgcgttgg caaagaggcc aaggacgcag atggcaacaa 4680  
acccaaggag cctgggataa cggaggcttc atcggagatg atatcaccaa acatgttgct 4740  
ggtgattata ataccatttta ggtgggttgg gttcttaactt aggtcatgg cggcagaatc 4800  
aatcaatttga tggtaactt tcaatgttgg gaattcgttc ttgtatgggg cttccacagt 4860

© 2002 Blackwell Science Ltd

ttttctccat aatcttgaag aggccaaaac attagctta tccaaggacc aaataggcaa 4920  
tggtggctca tgtttaggg ccatgaaagc ggccattctt gtgattctt gcacttctgg 4980  
aacggtgtat tgttcaactat cccaagcgac accatcacca tgcgttctt ttctcttacc 5040  
aaagtaaata cctcccacta attctctaac aacaacgaag tcagtacctt tagcaaattg 5100  
tggcttgatt ggagataagt ctaaaagaga gtcggatgca aagttacatg gtcttaagtt 5160  
ggcgtacaat tgaagttctt tacggatttt tagtaaacct tgcgttctt taacactacc 5220  
ggtaccccat ttaggaccac ccacagcacc taacaaaacg gcatcagcct tcttggaggc 5280  
ttccagcgcc tcacatcgaa gtggaacacc tgcgttctt atgcagcac caccaattaa 5340  
atgattttcg aaatcgaact tgacattgga acgaacatca gaaatagctt taagaacott 5400  
aatggcttcg gctgtgattt cttgaccaac gtggcacctt ggcaaaaacga cgatcttctt 5460  
aggggcagac attacaatgg tatatcctt aatatataat aaaaaaaaaa aaaaaaaaaa 5520  
aaaaaaaaaa atgcagcttc tcaatgatat tcgaatacgc tttgaggaga tacagcctaa 5580  
tatccgacaa actgtttac agattacga tcgtacttgt tacccatcat tgaattttga 5640  
acatccgaac ctgggagttt tccctgaaac agatagtata tttgaacctg tataataata 5700  
tatagtctag cgcttacgg aagacaatgt atgtatttcg gttcctggag aaactattgc 5760  
atctattgca taggtaatct tgcacgtcgc atcccggtt cattttctgc gtttccatct 5820  
tgcaactcaa tagcatatct ttgttaacga agcatctgtg cttcattttg tagaacaaaa 5880  
atgcaacgca agagcgctaa ttttcaaac aaagaatctg agctgcattt ttacagaaca 5940  
gaaatgcaac gcgaaagcgc tattttacca acgaagaatc tgcgttcat tttgtaaaa 6000  
caaaaatgca acgcgagagc gctaattttt caaacaaga atctgagctg catttttaca 6060  
gaacagaaat gcaacgca gacgcgtat tcccaacaaa gaatctatac ttcttttttg 6120  
ttctacaaaaa atgcacccccg agagcgctat ttttctaaca aagcatcttta gattactttt 6180  
tttctccttt gtgcgtctta taatgcagtc tttgataac ttttgcact gtaggtccgt 6240  
taaggttaga agaaggctac tttgggtgtt atttctctt ccataaaaaa agcctgactc 6300  
cacttccgc gtttactgat tactagcgaa gctgcgggtg catttttca agataaaggc 6360  
atccccgatt atattctata ccgatgtgaa ttgcgcatac tttgtgaaca gaaagtgata 6420  
gcgttgatga ttcttcattt gtcagaaaaat tatgaacggt ttcttctatt ttgtctctat 6480  
atactacgta taggaaatgt ttacatttc gtattgtttt cgattcactc tatgaatagt 6540  
tcttactaca attttttgtt ctaaagagta atactagaga taaacataaa aaatgttagag 6600

DRAFT - NOT FOR DISTRIBUTION

gtcgagttt aatgcgtttt 6660  
agcacagaga tatatacgaa agagatactt ttgagcaatg tttgtggaaag cggtattcgc 6720  
aatatTTAG tagctcgta cagtccggtg cgTTTGGT ttttgaaag tgcgtttca 6780  
gagcgctttt ggTTTCAAAG agcgctctga agttcctata ctttcttagag aataggaact 6840  
tcggaatagg aacttcaaAG cgTTTCCGAA aacgagcgct tccgaaaatg caacgcgagc 6900  
tgcgcacata cagctcactg ttcacgtcgc acctataatct gcgtgttgc ttttatata 6960  
tatacatgag aagaacggca tagtgcgtgt ttatgctaa atgcgtactt atatgcgtct 7020  
atttatgttag gatgaaaggt agtctagtagc ctcctgtat attatcccatt tccatgcggg 7080  
gtatcgatag ctTCCTTCAG cactaccctt tagctttctt atatgctgcc actcctcaat 7140  
tggatttagtc tcATCCTTCA atgctatcat ttccTTGTAT attggatcat atgcatacgta 7200  
ccgagaaact agtgcgaagt agtgcgttatttgcgtt atctgatgag tatacgTTGT 7260  
cctggccacg gcagaagcac gcttacgcgtt ccaatttccc acaacattag tcaactccgt 7320  
taggccccttc attgaaagaa atgaggtcat caaatgtctt ccaatgtgag attttggggcc 7380  
atTTTTATA gcaaagattt aataaggcgc atTTTCTTC aaagctttat tgtacgatct 7440  
gactaagtta tctttataata attggatttc ctgtttatttgc ttgaagaat tgccggcct 7500  
atttactcgat ttttagactg gttcagaatt cctaaaaat tcatccaaat atacaagtgg 7560  
atcgatgata agctgtcaaa catgagaattt cttgaagacg aaagggcctc gtgatacgcc 7620  
tattttata ggttaatgtc atgataataa tggTTCTTA gacgtcaggt ggcactttc 7680  
ggggaaatgt gcgcgaaacc cctatttgcattttctt aatacattca aatatgtatc 7740  
cgctcatgag acaataaccc tgataaaatgc ttcaataata ttgaaaaagg aagagtatga 7800  
gtattcaaca ttcccggtc gcccTTATTC cctttttgc ggcattttgc cttcctgttt 7860  
ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttt ggtgcacgag 7920  
tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagttt cccccgaaag 7980  
aacgtttcc aatgatgagc acttttaaag ttctgctatg tggcgcgta ttatccgtt 8040  
ttgacgcccgg gcaagagcaa ctgggtcgcc gcatacacta ttctcagaat gacttgggtt 8100  
agtactcacc agtcacagaa aagcatctt cggatggcat gacagtaaga gaattatgca 8160  
gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca acgatcgag 8220  
gaccgaagga gctaaccgct ttttgcaca acatggggga tcatgtact cgccttgatc 8280  
gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg 8340

cagcaatggc aacaacgttg cgcaaactat taactggcg aactacttact ctagcttccc 8400  
ggcaacaatt aatagactgg atggaggcg aataaagtgc aggaccactt ctgcgatcg 8460  
cccttcggc tggctggttt attgctgata aatctggagc cggtgagcgt gggatcg 8520  
gtatcattgc agcaactgggg ccagatggta agccctcccg tatacgatgtt atcacacga 8580  
cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac 8640  
tgattaagca ttggtaactg tcagaccaag tttactcata tatactttatg attgattaa 8700  
aacttcattt ttaatttaaa agatctagg tgaagatcct ttttataat ctcatgacca 8760  
aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag 8820  
gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaaca aaaaaaccac 8880  
cgctaccagc ggtggttgt ttgccggatc aagagctacc aactctttt ccgaaggtaa 8940  
ctggcttcag cagagcgcag ataccaaata ctgtccttct agtgtagccg tagttaggc 9000  
accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaatac ctgttaccag 9060  
tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac 9120  
cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc agttggagc 9180  
gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc 9240  
ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcggaaaca ggagagcgc 9300  
cgagggagct tccaggggaa aacgccttgtt atctttatag tccctgtcggg tttcgccacc 9360  
tctgacttga gcgtcgattt ttgtgatgtc cgtcaggggg gcggagccta tggaaaaacg 9420  
ccagcaacgc ggcctttta cggttccctgg ccttttgtg gcctttgct cacatgttct 9480  
ttcctgcgtt atccccgtat tctgtggata accgttattac cgcctttgag tgagctgata 9540  
ccgctcgccg cagccgaacg accgagcgc a gcgagtcgt gaggcaggaa gcggaaagagc 9600  
gcctgatgcg gtatttctc cttacgcattc tttgcgttat ttcacaccgc atatggtgca 9660  
ctctcagtac aatctgtctt gatgcccatc agttaagcca gtatacactc cgctatcgct 9720  
acgtgactgg gtcatggctg cggccggaca cccgccaaca cccgctgacg cgcctgacg 9780  
ggcttgtctg ctccggcat cggcttacag acaagctgtg accgtctccg ggagctgcat 9840  
gtgtcagagg ttttccacgtt catcaccgaa acgcgcgagg cagctgcgtt aaagctcatc 9900  
agcgtggctg tgaagcgatt cacagatgtc tgcctgttca tccgcgttcca gtcgttgag 9960  
tttctccaga agcgttaatg tctggcttct gataaagcgg gccatgtttaa gggcggtttt 10020  
ttcctgtttt gtcactgtatc ctcctgtta agggggattt ctgttcatgg ggtaatgtat 10080

accgatgaaa cgagagagga tgctcacat acgggtact gatgatgaac atgcccggtt 10140  
actggAACGT tgtgagggtt aacaactggc ggtatggatg cggcgggacc agagaaaaat 10200  
caactcagggt caatgccagc gcttcgttaa tacagatgtt ggtgttccac aggtagcca 10260  
gcagcatcct gcgtatgcaga tccggaaacat aatggtgcag ggcgtgact tcccggttcc 10320  
cagactttac gaaacacgga aaccgaagac cattcatgtt gttgctcagg tcgcagacgt 10380  
tttgcagcag cagtgcgttc acgttcgttc gcgtatcggt gattcattct gctaaccagt 10440  
aaggcaaccc cgccagccta gccgggtcct caacgacagg agcacgatca tggcaccgg 10500  
tggccaggac ccaacgctgc ccgagatgcg ccgcgtgcgg ctgctggaga tggcggacgc 10560  
gatggatatg ttctgccaag gttgggttg cgcattcaca gtttccgca agaattgatt 10620  
ggctccaatt cttggagtgg tgaatccgtt agcgaggtgc cgccgggttc cattcaggtc 10680  
gaggtggccc ggctccatgc accgcgacgc aacgcgggaa ggcagacaag gtatagggcg 10740  
gcccctacaa tccatgcca ccgttccat gtgctgcgg aggccgcata aatgcgcgtg 10800  
acgatcagcg gtccaatgat cgaagttagg ctggtaagag ccgcgaggaa tccttgaagc 10860  
tgtccctgat ggtcgcatc tacctgcctg gacagcatgg cctgcaacgc gggcatcccg 10920  
atgcccgg aagcgagaag aatcataatg gggaggcca tccagcctcg cgtcgcaac 10980  
gccagcaaga cgtagcccg cgcgtcgccg gccatgcgg cgataatggc ctgttctcg 11040  
ccgaaacgtt tggtggcgaa accagtgcacg aaggcttgag cgagggcgtg caagattccg 11100  
aataccgcaa gcgcacaggcc gatcatgtc gcgcgtccacg gaaagcggtc ctgcggaaa 11160  
atgacccaga gcgcgtccgg cacctgtcct acgagttgca tgataaagaa gacagtcata 11220  
agtgcggcga cgatagtcat gccccgcgcc caccggagg agctgactgg gttgaaggct 11280  
ctcaaggccca tcggtcgagg atccttcaat atgcgcacat acgctgttat gttcaaggtc 11340  
ccttcgttta agaacgaaag cggtcttcct tttgagggt gttcaagtt gttcaaatct 11400  
atcaaatttg caaatccccca gtctgtatct agagcggtga atcggtgatg cgatttggta 11460  
ataaaattga tggtgtcacc attaccaggc ctagatatac caatggcaaa ctgagcacaa 11520  
caataccagt ccggatcaac tggcaccatc tctccgttag tctcatctaa ttttcttcc 11580  
ggatgagggtt ccagatatac cgcaacaccc ttatattgtt ttccctgagg gaataataga 11640  
atgtcccatt cgaaatcacc aattctaaac ctggcgaat tggatccgg gtttggtaac 11700  
tcgttccagt caggaatgtt ccacgtgaag ctatcttcca gcaaagtctc cacttcttca 11760  
tcaaatttgtt gagaataactc ccaatgtct tatctatggg acttccggaa aaccacagtac 11820

DRAFT 2000

cgataacttcc caattcgtct tcagagctca ttgtttgttt gaagagacta atcaaagaat 11880  
cgaaaaattaa tatcttaact gatagttga tcaaaggggc aaaacgtagg 11940  
ggcaaacaaa cggaaaaatc gtttctcaa ttttctgtat ccaagaactc taaccagtct 12000  
tatctaaaaa ttgccttatg atccgtctct ccgggtacag cctgtgtaac tgattaatcc 12060  
tgcccttcta atcaccatc taatgttta attaaggat tttgtttca ttaacggctt 12120  
tcgctcataa aaatgttatg acgtttgcc cgccaggccc aaaccatcca ctccacgaga 12180  
ctgatctcct ctgccccac accgggcattc tccacttat aagttggaga aataagagaa 12240  
tttcagattt agagaatgaa aaaaaaaaaac ctttagttca taggtccatt ctcttagcgc 12300  
aactacagag aacagggca caaacaggca aaaaacgggc acaacctcaa tggagtgtat 12360  
caacctgcct ggagtaaatg atgacacaag gcaattgacc cacgcatgtat tctatctcat 12420  
tttcttacac cttcttattac cttctgtct ctctgatttg gaaaaagctg aaaaaaaaaagg 12480  
ttgaaaccag ttccctgaaa ttattccct acttgactaa taagtatata aagacggtag 12540  
gtattgattt taattctgtat aatctatttc tttaacttct taaattctac ttttatagtt 12600  
agtctttttt ttagtttaa aacaccaaga acttagttc gaataaacac acataaaca 12660  
acaagcttac aaaacaaa atg gct gca tat gca gct cag ggc tat aag gtg 12711  
Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val  
1 5 10  
ctat gta ctc aac ccc tct gtt gct gca aca ctg ggc ttt ggt gct tac 12759  
Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr  
15 20 25  
atg tcc aag gct cat ggg atc gat cct aac atc agg acc ggg gtg aga 12807  
Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg  
30 35 40  
aca att acc act ggc agc ccc atc acg tac tcc acc tac ggc aag ttc 12855  
Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe  
45 50 55  
ctt gcc gac ggc ggg tgc tcg ggg ggc gct tat gac ata ata att tgt 12903  
Leu Ala Asp Gly Gly Cys Ser Gly Ala Tyr Asp Ile Ile Ile Cys  
60 65 70 75  
gac gag tgc cac tcc acg gat gcc aca tcc atc ttg ggc att ggc act 12951  
Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr  
80 85 90  
gtc ctt gac caa gca gag act ggg ggg ggc aga ctg gtt gtg ctc gcc 12999  
Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala  
95 100 105  
acc gcc acc cct ccg ggc tcc gtc act gtg ccc cat ccc aac atc gag 13047

DRAFT

Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn Ile Glu				
110	115	120		
gag gtt gct ctg tcc acc acc gga gag atc cct ttt tac ggc aag gct				13095
Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala				
125	130	135		
atc ccc ctc gaa gta atc aag ggg ggg aga cat ctc atc ttc tgt cat				13143
Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu Ile Phe Cys His				
140	145	150	155	
tca aag aag aag tgc gac gaa ctc gcc gca aag ctg gtc gca ttg ggc				13191
Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Ala Leu Gly				
160	165	170		
atc aat gcc gtg gcc tac tac cgc ggt ctt gac gtg tcc gtc atc ccg				13239
Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile Pro				
175	180	185		
acc agc ggc gat gtt gtc gtc gca acc gat gcc ctc atg acc ggc				13287
Thr Ser Gly Asp Val Val Val Ala Thr Asp Ala Leu Met Thr Gly				
190	195	200		
tat acc ggc gac ttc gac tcg gtg ata gac tgc aat acg tgt gtc acc				13335
Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys Val Thr				
205	210	215		
cag aca gtc gat ttc agc ctt gac cct acc ttc acc att gag aca atc				13383
Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr Ile				
220	225	230	235	
acg ctc ccc caa gat gct gtc tcc cgc act caa cgt cggt ggc agg act				13431
Thr Leu Pro Gln Asp Ala Val Ser Arg Thr Gln Arg Arg Gly Arg Thr				
240	245	250		
ggc agg ggg aag cca ggc atc tac aga ttt gtg gca ccg ggg gag cgc				13479
Gly Arg Gly Lys Pro Gly Ile Tyr Arg Phe Val Ala Pro Gly Glu Arg				
255	260	265		
ccc tcc ggc atg ttc gac tcg tcc gtc ctc tgt gag tgc tat gac gca				13527
Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp Ala				
270	275	280		
ggc tgt gct tgg tat gag ctc acg ccc gcc gag act aca gtt agg cta				13575
Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Thr Val Arg Leu				
285	290	295		
cga gcg tac atg aac acc ccg ggg ctt ccc gtg tgc cag gac cat ctt				13623
Arg Ala Tyr Met Asn Thr Pro Gly Leu Pro Val Cys Gln Asp His Leu				
300	305	310	315	
gaa ttt tgg gag ggc gtc ttt aca ggc ctc act cat ata gat gcc cac				13671
Glu Phe Trp Glu Gly Val Phe Thr Gly Leu Thr His Ile Asp Ala His				
320	325	330		
ttt cta tcc cag aca aag cag agt ggg gag aac ctt cct tac ctg gta				13719
Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Leu Pro Tyr Leu Val				

DRAFT

335	340	345			
gcg tac caa gcc acc gtg tgc gct agg gct caa gcc cct ccc cca tcg Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser			13767		
350	355	360			
tgg gac cag atg tgg aag tgt ttg att cgc ctc aag ccc acc ctc cat Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His			13815		
365	370	375			
ggg cca aca ccc ctg cta tac aga ctg ggc gct gtt cag aat gaa atc Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Ile			13863		
380	385	390	395		
acc ctg acg cac cca gtc acc aaa tac atc atg aca tgc atg tcg gcc Thr Leu Thr His Pro Val Thr Lys Tyr Ile Met Thr Cys Met Ser Ala	400	405	410	13911	
gac ctg gag gtc acg agc acc tgg gtg ctc gtt ggc ggc gtc ctg Asp Leu Glu Val Val Thr Ser Thr Trp Val Leu Val Gly Gly Val Leu	415	420	425	13959	
430	435	440			
gct gct ttg gcc gcg tat tgc ctg tca aca ggc tgc gtg gtc ata gtg Ala Ala Leu Ala Ala Tyr Cys Leu Ser Thr Gly Cys Val Val Ile Val			14007		
ggc agg gtc gtc ttg tcc ggg aag ccg gca atc ata cct gac agg gaa Gly Arg Val Val Leu Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu	445	450	455	14055	
gtc ctc tac cga gag ttc gat gag atg gaa gag tgc tct cag cac tta Val Leu Tyr Arg Glu Asp Glu Met Glu Glu Cys Ser Gln His Leu	460	465	470	475	14103
ccg tac atc gag caa ggg atg atg ctc gcc gag cag ttc aag cag aag Pro Tyr Ile Glu Gln Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys	480	485	490	14151	
gcc ctc ggc ctc ctg cag acc gcg tcc cgt cag gca gag gtt atc gcc Ala Leu Gly Leu Leu Gln Thr Ala Ser Arg Gln Ala Glu Val Ile Ala	495	500	505	14199	
cct gct gtc cag acc aac tgg caa aaa ctc gag acc ttc tgg gcg aag Pro Ala Val Gln Thr Asn Trp Gln Lys Leu Glu Thr Phe Trp Ala Lys	510	515	520	14247	
cat atg tgg aac ttc atc agt ggg ata caa tac ttg gcg ggc ttg tca His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala Gly Leu Ser	525	530	535	14295	
acg ctg cct ggt aac ccc gcc att gct tca ttg atg gct ttt aca gct Thr Leu Pro Gly Asn Pro Ala Ile Ala Ser Leu Met Ala Phe Thr Ala	540	545	550	555	14343
gct gtc acc agc cca cta acc act agc caa acc ctc ctc ttc aac ata Ala Val Thr Ser Pro Leu Thr Thr Ser Gln Thr Leu Leu Phe Asn Ile	560	565	570	14391	

100

ttg ggg ggg tgg gtg gct gcc cag ctc gcc gcc ccc ggt gcc gct act Leu Gly Gly Trp Val Ala Ala Gln Leu Ala Ala Pro Gly Ala Ala Thr 575 580 585	14439
gcc ttt gtg ggc gct ggc tta gct ggc gcc gcc atc ggc agt gtt gga Ala Phe Val Gly Ala Gly Leu Ala Gly Ala Ile Gly Ser Val Gly 590 595 600	14487
ctg ggg aag gtc ctc ata gac atc ctt gca ggg tat ggc gcg ggc gtg Leu Gly Lys Val Leu Ile Asp Ile Leu Ala Gly Tyr Gly Ala Gly Val 605 610 615	14535
gcg gga gct ctt gtg gca ttc aag atc atg agc ggt gag gtc gcc tcc Ala Gly Ala Leu Val Ala Phe Lys Ile Met Ser Gly Glu Val Pro Ser 620 625 630 635	14583
acg gag gac ctg gtc aat cta ctg ccc gcc atc ctc tcg ccc gga gcc Thr Glu Asp Leu Val Asn Leu Leu Pro Ala Ile Leu Ser Pro Gly Ala 640 645 650	14631
ctc gta gtc ggc gtg gtc tgt gca gca ata ctg cgc cgcc cac gtt ggc Leu Val Val Gly Val Val Cys Ala Ala Ile Leu Arg Arg His Val Gly 655 660 665	14679
ccg ggc gag ggg gca gtg cag tgg atg aac cggtg ataa gcc ttc gcc Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile Ala Phe Ala 670 675 680	14727
tcc cgg ggg aac cat gtt tcc ccc acg cac tac gtg ccg gag agc gat Ser Arg Gly Asn His Val Ser Pro Thr His Tyr Val Pro Glu Ser Asp 685 690 695	14775
gca gct gcc cgc gtc act gcc ata ctc agc agc ctc act gta acc cag Ala Ala Ala Arg Val Thr Ala Ile Leu Ser Ser Leu Thr Val Thr Gln 700 705 710 715	14823
ctc ctg agg cga ctg cac cag tgg ata agc ccgtg gag tgt acc act cca Leu Leu Arg Arg Leu His Gln Trp Ile Ser Ser Glu Cys Thr Thr Pro 720 725 730	14871
tgc tcc ggt tcc tgg cta agg gac atc tgg gac tgg ata tgc gag gtg Cys Ser Gly Ser Trp Leu Arg Asp Ile Trp Asp Trp Ile Cys Glu Val 735 740 745	14919
ttg agc gac ttt aag acc tgg cta aaa gct aag ctc atg cca cag ctg Leu Ser Asp Phe Lys Thr Trp Leu Lys Ala Lys Leu Met Pro Gln Leu 750 755 760	14967
cct ggg atc ccc ttt gtg tcc tgc cag cgc ggg tat aag ggg gtc tgg Pro Gly Ile Pro Phe Val Ser Cys Gln Arg Gly Tyr Lys Gly Val Trp 765 770 775	15015
cga ggg gac ggc atc atg cac act cgc tgc cac tgt gga gct gag atc Arg Gly Asp Gly Ile Met His Thr Arg Cys His Cys Gly Ala Glu Ile 780 785 790 795	15063
act gga cat gtc aaa aac ggg acg atg agg atc gtc ggt cct agg acc	15111

Thr Gly His Val Lys Asn Gly Thr Met Arg Ile Val Gly Pro Arg Thr  
 800 805 810

tgc agg aac atg tgg agt ggg acc ttc ccc att aat gcc tac acc acg 15159  
 Cys Arg Asn Met Trp Ser Gly Thr Phe Pro Ile Asn Ala Tyr Thr Thr  
 815 820 825

ggc ccc tgt acc ccc ctt cct gcg ccg aac tac acg ttc gcg cta tgg 15207  
 Gly Pro Cys Thr Pro Leu Pro Ala Pro Asn Tyr Thr Phe Ala Leu Trp  
 830 835 840

agg gtg tct gca gag gaa tac gtg gag ata agg cag gtg ggg gac ttc 15255  
 Arg Val Ser Ala Glu Glu Tyr Val Glu Ile Arg Gln Val Gly Asp Phe  
 845 850 855

cac tac gtg acg ggt atg act act gac aat ctt aaa tgc ccg tgc cag 15303  
 His Tyr Val Thr Gly Met Thr Asp Asn Leu Lys Cys Pro Cys Gln  
 860 865 870 875

gtc cca tcg ccc gaa ttt ttc aca gaa ttg gac ggg gtg cgc cta cat 15351  
 Val Pro Ser Pro Glu Phe Phe Thr Glu Leu Asp Gly Val Arg Leu His  
 880 885 890

agg ttt gcg ccc ccc tgc aag ccc ttg ctg cg<sup>g</sup> gag gaa tca ttc 15399  
 Arg Phe Ala Pro Pro Cys Lys Pro Leu Leu Arg Glu Glu Val Ser Phe  
 895 900 905

aga gta gga ctc cac gaa tac ccg gta ggg tcg caa tta cct tgc gag 15447  
 Arg Val Gly Leu His Glu Tyr Pro Val Gly Ser Gln Leu Pro Cys Glu  
 910 915 920

ccc gaa ccg gac gtg gcc gtg ttg acg tcc atg ctc act gat ccc tcc 15495  
 Pro Glu Pro Asp Val Ala Val Leu Thr Ser Met Leu Thr Asp Pro Ser  
 925 930 935

cat ata aca gca gag gcg gcc ggg cga agg ttg gcg agg gga tca ccc 15543  
 His Ile Thr Ala Glu Ala Gly Arg Arg Leu Ala Arg Gly Ser Pro  
 940 945 950 955

ccc tct gtg gcc agc tcc tcg gct agc cag cta tcc gct cca tct ctc 15591  
 Pro Ser Val Ala Ser Ser Ala Ser Gln Leu Ser Ala Pro Ser Leu  
 960 965 970

aag gca act tgc acc gct aac cat gac tcc cct gat gct gag ctc ata 15639  
 Lys Ala Thr Cys Thr Ala Asn His Asp Ser Pro Asp Ala Glu Leu Ile  
 975 980 985

gag gcc aac ctc cta tgg agg cag gag atg ggc ggc aac atc acc agg 15687  
 Glu Ala Asn Leu Leu Trp Arg Gln Glu Met Gly Gly Asn Ile Thr Arg  
 990 995 1000

gtt gag tca gaa aac aaa gtg gtg att ctg gac tcc ttc gat ccg ctt 15735  
 Val Glu Ser Glu Asn Lys Val Val Ile Leu Asp Ser Phe Asp Pro Leu  
 1005 1010 1015

gtg gcg gag gag gac gag cgg gag atc tcc gta ccc gca gaa atc ctg 15783  
 Val Ala Glu Glu Asp Glu Arg Glu Ile Ser Val Pro Ala Glu Ile Leu

1020                  1025                  1030                  1035

cg<sup>g</sup> a<sup>a</sup> g<sup>t</sup> t<sup>c</sup> c<sup>g</sup> a<sup>g</sup> t<sup>t</sup> c<sup>g</sup> c<sup>g</sup> c<sup>g</sup> c<sup>g</sup> c<sup>g</sup> c<sup>g</sup> c<sup>g</sup>  
 Arg Lys Ser Arg Arg Phe Ala Gln Ala Leu Pro Val Trp Ala Arg Pro  
 1040                  1045                  1050

15831

g<sup>a</sup>c t<sup>a</sup>t a<sup>a</sup>c c<sup>cc</sup> c<sup>cg</sup> c<sup>t</sup>a g<sup>t</sup>g g<sup>a</sup>g a<sup>c</sup>g t<sup>gg</sup> a<sup>aa</sup> a<sup>a</sup>g c<sup>cc</sup> g<sup>a</sup>c t<sup>a</sup>c g<sup>a</sup>a  
 Asp Tyr Asn Pro Pro Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr Glu  
 1055                  1060                  1065

15879

c<sup>cc</sup> c<sup>c</sup>t g<sup>t</sup>g g<sup>t</sup>c c<sup>a</sup>t g<sup>g</sup>c t<sup>g</sup>c c<sup>cc</sup> c<sup>tt</sup> c<sup>ca</sup> a<sup>a</sup>g t<sup>c</sup>c c<sup>c</sup>t c<sup>c</sup>t  
 Pro Pro Val Val His Gly Cys Pro Leu Pro Pro Pro Lys Ser Pro Pro  
 1070                  1075                  1080

15927

g<sup>t</sup>g c<sup>c</sup>t c<sup>cc</sup> c<sup>c</sup>t c<sup>gg</sup> a<sup>a</sup>g a<sup>a</sup>g c<sup>gg</sup> a<sup>c</sup>g g<sup>t</sup>g g<sup>t</sup>c c<sup>tc</sup> a<sup>c</sup>t g<sup>a</sup>a t<sup>c</sup>a a<sup>cc</sup>  
 Val Pro Pro Pro Arg Lys Lys Arg Thr Val Val Leu Thr Glu Ser Thr  
 1085                  1090                  1095

15975

c<sup>t</sup>a t<sup>c</sup>t a<sup>c</sup>t g<sup>cc</sup> t<sup>tg</sup> g<sup>cc</sup> g<sup>a</sup>g c<sup>t</sup>c g<sup>cc</sup> a<sup>cc</sup> a<sup>ga</sup> a<sup>gc</sup> t<sup>tt</sup> g<sup>gc</sup> a<sup>gc</sup> t<sup>c</sup>c  
 Leu Ser Thr Ala Leu Ala Glu Leu Ala Thr Arg Ser Phe Gly Ser Ser  
 1100                  1105                  1110                  1115

16023

t<sup>c</sup>a a<sup>c</sup>t t<sup>c</sup>c g<sup>gc</sup> a<sup>t</sup>t a<sup>c</sup>g g<sup>gc</sup> g<sup>a</sup>at a<sup>c</sup>t a<sup>c</sup>a a<sup>c</sup>a t<sup>c</sup>c t<sup>c</sup>t g<sup>a</sup>g c<sup>cc</sup>  
 Ser Thr Ser Gly Ile Thr Gly Asp Asn Thr Thr Ser Ser Glu Pro  
 1120                  1125                  1130

16071

g<sup>cc</sup> c<sup>c</sup>t t<sup>c</sup>t g<sup>gc</sup> t<sup>gc</sup> c<sup>cc</sup> c<sup>cc</sup> g<sup>a</sup>c t<sup>c</sup>c g<sup>c</sup>t g<sup>a</sup>g t<sup>c</sup>c t<sup>a</sup>t t<sup>c</sup>c t<sup>c</sup>c  
 Ala Pro Ser Gly Cys Pro Pro Asp Ser Asp Ala Glu Ser Tyr Ser Ser  
 1135                  1140                  1145

16119

a<sup>t</sup>g c<sup>cc</sup> c<sup>cc</sup> c<sup>t</sup>g g<sup>a</sup>g g<sup>gg</sup> g<sup>a</sup>g c<sup>c</sup>t g<sup>gg</sup> g<sup>a</sup>t c<sup>cg</sup> g<sup>a</sup>t c<sup>tt</sup> a<sup>gc</sup> g<sup>ac</sup> g<sup>gg</sup>  
 Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu Ser Asp Gly  
 1150                  1155                  1160

16167

t<sup>c</sup>a t<sup>gg</sup> t<sup>c</sup>a a<sup>c</sup>g g<sup>t</sup>c a<sup>gt</sup> a<sup>gt</sup> g<sup>a</sup>g g<sup>c</sup>c a<sup>a</sup>c g<sup>gg</sup> g<sup>a</sup>g g<sup>a</sup>t g<sup>t</sup>c g<sup>t</sup>g t<sup>gc</sup>  
 Ser Trp Ser Thr Val Ser Ser Glu Ala Asn Ala Glu Asp Val Val Cys  
 1165                  1170                  1175

16215

t<sup>gc</sup> t<sup>c</sup>a a<sup>t</sup>g t<sup>c</sup>t t<sup>gg</sup> a<sup>c</sup>a g<sup>gc</sup> g<sup>cc</sup> g<sup>ca</sup> c<sup>tc</sup> g<sup>t</sup>c a<sup>cc</sup> c<sup>cg</sup> t<sup>gc</sup> g<sup>cc</sup>  
 Cys Ser Met Ser Tyr Ser Trp Thr Gly Ala Leu Val Thr Pro Cys Ala  
 1180                  1185                  1190                  1195

16263

g<sup>c</sup>g g<sup>a</sup>a g<sup>a</sup>a c<sup>a</sup>g a<sup>aa</sup> c<sup>t</sup>g c<sup>cc</sup> a<sup>t</sup>c a<sup>a</sup>t g<sup>c</sup>a c<sup>t</sup>a a<sup>a</sup>c t<sup>c</sup>g t<sup>tg</sup> c<sup>t</sup>a  
 Ala Glu Glu Gln Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu  
 1200                  1205                  1210

16311

c<sup>gt</sup> c<sup>a</sup>c c<sup>a</sup>c a<sup>a</sup>t t<sup>tg</sup> g<sup>t</sup>g t<sup>a</sup>t t<sup>c</sup>c a<sup>c</sup>c a<sup>c</sup>c t<sup>c</sup>a c<sup>gc</sup> a<sup>gt</sup> g<sup>c</sup>t t<sup>gc</sup> c<sup>aa</sup>  
 Arg His His Asn Leu Val Tyr Ser Thr Ser Arg Ser Ala Cys Gln  
 1215                  1220                  1225

16359

a<sup>gg</sup> c<sup>a</sup>g a<sup>aa</sup> g<sup>t</sup>c a<sup>c</sup>a t<sup>tt</sup> g<sup>a</sup>c a<sup>ga</sup> c<sup>aa</sup> g<sup>t</sup>t c<sup>tg</sup> g<sup>a</sup>c a<sup>gc</sup> a<sup>c</sup>t  
 Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Ser His  
 1230                  1235                  1240

16407

t<sup>a</sup>c c<sup>a</sup>g g<sup>t</sup>a c<sup>t</sup>c a<sup>a</sup>g g<sup>a</sup>g g<sup>t</sup>t a<sup>aa</sup> g<sup>c</sup>a g<sup>c</sup>g g<sup>c</sup>g t<sup>c</sup>a a<sup>aa</sup> g<sup>t</sup>g a<sup>a</sup>g  
 Tyr Gln Asp Val Leu Lys Glu Val Lys Ala Ala Ser Lys Val Lys  
 1245                  1250                  1255

16455

gct aac ttg cta tcc gta gag gaa gct tgc agc ctg acg ccc cca cac 16503  
 Ala Asn Leu Leu Ser Val Glu Glu Ala Cys Ser Leu Thr Pro Pro His  
 1260 1265 1270 1275

tca gcc aaa tcc aag ttt ggt tat ggg gca aaa gac gtc cgt tgc cat 16551  
 Ser Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Cys His  
 1280 1285 1290

gcc aga aag gcc gta acc cac atc aac tcc gtg tgg aaa gac ctt ctg 16599  
 Ala Arg Lys Ala Val Thr His Ile Asn Ser Val Trp Lys Asp Leu Leu  
 1295 1300 1305

gaa gac aat gta aca cca ata gac act acc atc atg gct aag aac gag 16647  
 Glu Asp Asn Val Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu  
 1310 1315 1320

gtt ttc tgc gtt cag cct gag aag ggg ggt cgt aag cca gct cgt ctc 16695  
 Val Phe Cys Val Gln Pro Glu Lys Gly Arg Lys Pro Ala Arg Leu  
 1325 1330 1335

atc gtg ttc ccc gat ctg ggc gtg cgc gtg tgc gaa aag atg gct ttg 16743  
 Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu  
 1340 1345 1350 1355

tac gac gtg gtt aca aag ctc ccc ttg gcc gtg atg gga agc tcc tac 16791  
 Tyr Asp Val Val Thr Lys Leu Pro Leu Ala Val Met Gly Ser Ser Tyr  
 1360 1365 1370

gga ttc caa tac tca cca gga cag cgg gtt gaa ttc ctc gtg caa gcg 16839  
 Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Gln Ala  
 1375 1380 1385

tgg aag tcc aag aaa acc cca atg ggg ttc tcg tat gat acc cgc tgc 16887  
 Trp Lys Ser Lys Lys Thr Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys  
 1390 1395 1400

ttt gac tcc aca gtc act gag agc gac atc cgt acg gag gag gca atc 16935  
 Phe Asp Ser Thr Val Thr Glu Ser Asp Ile Arg Thr Glu Glu Ala Ile  
 1405 1410 1415

tac caa tgt tgt gac ctc gac ccc caa gcc cgc gtg gcc atc aag tcc 16983  
 Tyr Gln Cys Cys Asp Leu Asp Pro Gln Ala Arg Val Ala Ile Lys Ser  
 1420 1425 1430 1435

ctc acc gag agg ctt tat gtt ggg ggc ctt acc aat tca agg ggg 17031  
 Leu Thr Glu Arg Leu Tyr Val Gly Gly Pro Leu Thr Asn Ser Arg Gly  
 1440 1445 1450

gag aac tgc ggc tat cgc agg tgc cdc gcg agc ggc gta ctg aca act 17079  
 Glu Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr  
 1455 1460 1465

agc tgt ggt aac acc ctc act tgc tac atc aag gcc cgg gca gcc tgt 17127  
 Ser Cys Gly Asn Thr Leu Thr Cys Tyr Ile Lys Ala Arg Ala Ala Cys  
 1470 1475 1480

cga gcc gca ggg ctc cag gac tgc acc atg ctc gtg tgt ggc gac gac 17175

Arg Ala Ala Gly Leu Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp  
 1485 1490 1495

tta gtc gtt atc tgt gaa agc gcg ggg gtc cag gag gac gcg gcg agc  
 1500 1505 1510 1515

17223

Leu Val Val Ile Cys Glu Ser Ala Gly Val Gln Glu Asp Ala Ala Ser

ctg aga gcc ttc acg gag gct atg acc agg tac tcc gcc ccc cct ggg  
 1520 1525 1530

17271

Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly

gac ccc cca caa cca gaa tac gac ttg gag ctc ata aca tca tgc tcc  
 1535 1540 1545

17319

Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser

tcc aac gtg tca gtc gcc cac gac ggc gct gga aag agg gtc tac tac  
 1550 1555 1560

17367

Ser Asn Val Ser Val Ala His Asp Gly Ala Gly Lys Arg Val Tyr Tyr

ctc acc cgt gac cct aca acc ccc ctc gcg aga gct gcg tgg gag aca  
 1565 1570 1575

17415

Leu Thr Arg Asp Pro Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr

gca aga cac act cca gtc aat tcc tgg cta ggc aac ata atc atg ttt  
 1580 1585 1590 1595

17463

Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Phe

gcc ccc aca ctg tgg gcg agg atg ata ctg atg acc cat ttc ttt agc  
 1600 1605 1610

17511

Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser

gtc ctt ata gcc agg gac cag ctt gaa cag gcc ctc gat tgc gag atc  
 1615 1620 1625

17559

Val Leu Ile Ala Arg Asp Gln Leu Glu Gln Ala Leu Asp Cys Glu Ile

tac ggg gcc tgc tac tcc ata gaa cca ctg gat cta cct cca atc att  
 1630 1635 1640

17607

Tyr Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro Pro Ile Ile

caa aga ctc cat ggc ctc agc gca ttt tca ctc cac agt tac tct cca  
 1645 1650 1655

17655

Gln Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro

ggt gaa atc aat agg gtg gcc gca tcc ctc aga aaa ctt ggg gta ccg  
 1660 1665 1670 1675

17703

Gly Glu Ile Asn Arg Val Ala Ala Cys Leu Arg Lys Leu Gly Val Pro

ccc ttg cga gct tgg aga cac cgg gcc cgg agc gtc cgc gct agg ctt  
 1680 1685 1690

17751

Pro Leu Arg Ala Trp Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu

ctg gcc aga gga ggc agg gct gcc ata tgt ggc aag tac ctc ttc aac  
 1695 1700 1705

17799

Leu Ala Arg Gly Gly Arg Ala Ala Ile Cys Gly Lys Tyr Leu Phe Asn

tgg gca gta aga aca aag ctc aaa ctc act cca ata gcg gcc gct ggc  
 17847

Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Ala Ala Gly

1710            1715            1720            17895

cag ctg gac ttg tcc ggc tgg ttc acg gct ggc tac agc ggg gga gac  
 Gln Leu Asp Leu Ser Gly Trp Phe Thr Ala Gly Tyr Ser Gly Gly Asp  
 1725            1730            1735

att tat cac agc gtg tct cat gcc cgg ccc cgc tgg atc tgg ttt tgc  
 Ile Tyr His Ser Val Ser His Ala Arg Pro Arg Trp Ile Trp Phe Cys  
 1740            1745            1750            1755            17943

cta ctc ctg ctt gct gca ggg gta ggc atc tac ctc ctc ccc aac cga  
 Leu Leu Leu Leu Ala Ala Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg  
 1760            1765            1770            17991

atg agc acg aat cct aaa cct caa aga aag acc aaa cgt aac acc aac  
 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn  
 1775            1780            1785            18039

cgg cgg ccg cag gac gtc aag ttc cgg ggt ggc ggt cag atc gtt ggt  
 Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gln Ile Val Gly  
 1790            1795            1800            18087

gga gtt tac ttg ttg ccg cgc agg ggc cct aga ttg ggt gtc cgc gcg  
 Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala  
 1805            1810            1815            18135

acg aga aag act tcc gag cgg tcg caa cct cga ggt aga cgt cag cct  
 Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro  
 1820            1825            1830            1835            18183

atc ccc aag gct cgt cgg ccc gag ggc agg acc tgg gct cag ccc ggg  
 Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly  
 1840            1845            1850            18231

tac cct tgg ccc ctc tat ggc aat gag ggc tgc ggg tgg gcg gga tgg  
 Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp  
 1855            1860            1865            18279

ctc ctg tct ccc cgt ggc tct cgg cct agc tgg ggc ccc aca gac ccc  
 Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro  
 1870            1875            1880            18327

cgg cgt agg tcg cgc aat ttg ggt aag taatagtcga ctttggcccc  
 Arg Arg Arg Ser Arg Asn Leu Gly Lys            18374

1885            1890

actgtacttt tagctcgtac aaaataacaat atactttca tttctccgtac aacaacatgt 18434

tttcccatgt aatatccctt tctattttc gttccgttac caactttaca catactttat 18494

atagctattc acttctatac actaaaaaac taagacaatt ttaattttgc tgcctgccat 18554

atttcaattt gttataaatt cctataat tt atcctttag tagctaaaaa aagatgaatg 18614

tgaatcgaat cctaaagagaa ttggatctga tccacaggac ggggtgtggtc gccatgatcg 18674

cgttagtcgat agtggctcca agtagcgaag cgagcaggac tgggcggcgg ccaaagcgg 18734

cggacagtgc tccgagaacg ggtgcgcata gaaattgcata caacgcata agcgctagca 18794  
gcacgccata gtgactggcg atgctgtcgg aatggacgat atcccgcaag aggcccggca 18854  
gtaccggcat aaccaaggct atgcctacag catccagggt gacggtgccg aggatgacga 18914  
tgagcgcatt gtttagatttc atacacggtg cctgactgctg tttagcaattt aactgtgata 18974  
aactaccgca ttaaagcttt ttctttccaa tttttttttt ttcgtcatta taaaaatcat 19034  
tacgaccgag attcccggtt aataactgat ataattaaat tgaagctcta atttgtgagt 19094  
tttagtataca tgcatttact tataatacag ttttttagtt ttgctggccg catcttctca 19154  
aatatgcttc ccagcctgct tttctgtAAC gttcacccctc taccttagca tcccttcct 19214  
ttgcaaatacg tcctcttcca acaataataa tgtcagatcc tgttagagacc acatcatcca 19274  
cggttctata ctgttgaccc aatgcgtctc ccttgtcatc taaaacccaca ccgggtgtca 19334  
taatcaacca atcgtaacct tcatctcttc cacccatgtc tctttgagca ataaagccga 19394  
taacaaaatc tttgtcgctc ttgcataatgt caacagtacc cttagtatat cttccagtag 19454  
atagggagcc cttgcatgac aattctgcta acatcaaaag gcctctaggt tcctttgtta 19514  
cttcttcgtc cgccctgcttc aaaccgctaa caatacctgg gcccaccaca ccgtgtgcat 19574  
tcgtaatgtc tgcccattct gctattctgt atacacccgc agagtacctc aatttgactg 19634  
tattaccaat gtcagcaaat tttctgtctt cgaagagtaa aaaattgtac ttggcggata 19694  
atgccttttag cggttaact gtgccttcca tggaaaaatc agtcaagata tccacatgt 19754  
tttttagtaa acaaattttg ggacctaattt cttcaactaa ctccagtaat tccttggtgg 19814  
tacgaacatc caatgaagca cacaagttt tttgttttc gtgcataatc taaaatagct 19874  
tggcagcaac aggacttagga tgagtagcag cacgttcctt atatgtatct ttcgacatga 19934  
tttatcttcg tttcctgcaag gttttgttc tttgtcaggat ggttaagaat actgggcaat 19994  
ttcatgtttc ttcaacacta catatgcgtt tatataccaa tcttaagtctg tgctccttcc 20054  
ttcgttcttc cttctgttcg gagattaccg aatcaaaaaa atttcaagga aaccgaaatc 20114  
aaaaaaaaaga ataaaaaaaaa aatgatgaat tgaaaagctt atcgat 20160

<210> 13  
<211> 1892  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
pd.delta.NS3NS5.pj.core121

<400> 13  
Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro  
1 5 10 15

Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His  
20 25 30

Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly  
35 40 45

Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly  
50 55 60

Cys Ser Gly Gly Ala Tyr Asp Ile Ile Cys Asp Glu Cys His Ser  
65 70 75 80

Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Ala  
85 90 95

Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala Thr Ala Thr Pro Pro  
100 105 110

Gly Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser  
115 120 125

Thr Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val  
130 135 140

Ile Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Lys Lys Cys  
145 150 155 160

Asp Glu Leu Ala Ala Lys Leu Val Ala Leu Gly Ile Asn Ala Val Ala  
165 170 175

Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val  
180 185 190

Val Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe  
195 200 205

Asp Ser Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe  
210 215 220

Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr Ile Thr Leu Pro Gln Asp  
225 230 235 240

Ala Val Ser Arg Thr Gln Arg Arg Gly Arg Thr Gly Arg Gly Lys Pro  
245 250 255

Gly Ile Tyr Arg Phe Val Ala Pro Gly Glu Arg Pro Ser Gly Met Phe  
260 265 270

Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp Ala Gly Cys Ala Trp Tyr  
275 280 285

Glu Leu Thr Pro Ala Glu Thr Thr Val Arg Leu Arg Ala Tyr Met Asn  
290 295 300

Thr Pro Gly Leu Pro Val Cys Gln Asp His Leu Glu Phe Trp Glu Gly  
305 310 315 320

Val Phe Thr Gly Leu Thr His Ile Asp Ala His Phe Leu Ser Gln Thr  
325 330 335

Lys Gln Ser Gly Glu Asn Leu Pro Tyr Leu Val Ala Tyr Gln Ala Thr  
340 345 350

Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser Trp Asp Gln Met Trp  
355 360 365

Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His Gly Pro Thr Pro Leu  
370 375 380

Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Ile Thr Leu Thr His Pro  
385 390 395 400

Val Thr Lys Tyr Ile Met Thr Cys Met Ser Ala Asp Leu Glu Val Val  
405 410 415

Thr Ser Thr Trp Val Leu Val Gly Val Leu Ala Ala Leu Ala Ala  
420 425 430

Tyr Cys Leu Ser Thr Gly Cys Val Val Ile Val Gly Arg Val Val Leu  
435 440 445

Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu Val Leu Tyr Arg Glu  
450 455 460

Phe Asp Glu Met Glu Glu Cys Ser Gln His Leu Pro Tyr Ile Glu Gln  
465 470 475 480

Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys Ala Leu Gly Leu Leu  
485 490 495

Gln Thr Ala Ser Arg Gln Ala Glu Val Ile Ala Pro Ala Val Gln Thr  
500 505 510

Asn Trp Gln Lys Leu Glu Thr Phe Trp Ala Lys His Met Trp Asn Phe  
515 520 525

Ile Ser Gly Ile Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn  
530 535 540

Pro Ala Ile Ala Ser Leu Met Ala Phe Thr Ala Ala Val Thr Ser Pro  
545 550 555 560

Leu Thr Thr Ser Gln Thr Leu Leu Phe Asn Ile Leu Gly Gly Trp Val  
565 570 575

Ala Ala Gln Leu Ala Ala Pro Gly Ala Ala Thr Ala Phe Val Gly Ala  
580 585 590

Gly Leu Ala Gly Ala Ala Ile Gly Ser Val Gly Leu Gly Lys Val Leu  
595 600 605

Ile Asp Ile Leu Ala Gly Tyr Gly Ala Gly Val Ala Gly Ala Leu Val  
610 615 620

Ala Phe Lys Ile Met Ser Gly Glu Val Pro Ser Thr Glu Asp Leu Val  
625 630 635 640

Asn Leu Leu Pro Ala Ile Leu Ser Pro Gly Ala Leu Val Val Gly Val  
645 650 655

Val Cys Ala Ala Ile Leu Arg Arg His Val Gly Pro Gly Glu Gly Ala  
660 665 670

Val Gln Trp Met Asn Arg Leu Ile Ala Phe Ala Ser Arg Gly Asn His  
675 680 685

Val Ser Pro Thr His Tyr Val Pro Glu Ser Asp Ala Ala Ala Arg Val  
690 695 700

Thr Ala Ile Leu Ser Ser Leu Thr Val Thr Gln Leu Leu Arg Arg Leu  
705 710 715 720

His Gln Trp Ile Ser Ser Glu Cys Thr Thr Pro Cys Ser Gly Ser Trp  
725 730 735

Leu Arg Asp Ile Trp Asp Trp Ile Cys Glu Val Leu Ser Asp Phe Lys  
740 745 750

Thr Trp Leu Lys Ala Lys Leu Met Pro Gln Leu Pro Gly Ile Pro Phe  
755 760 765

Val Ser Cys Gln Arg Gly Tyr Lys Gly Val Trp Arg Gly Asp Gly Ile  
770 775 780

Met His Thr Arg Cys His Cys Gly Ala Glu Ile Thr Gly His Val Lys  
785 790 795 800

Asn Gly Thr Met Arg Ile Val Gly Pro Arg Thr Cys Arg Asn Met Trp  
805 810 815

Ser Gly Thr Phe Pro Ile Asn Ala Tyr Thr Thr Gly Pro Cys Thr Pro  
820 825 830

Leu Pro Ala Pro Asn Tyr Thr Phe Ala Leu Trp Arg Val Ser Ala Glu  
835 840 845

Glu Tyr Val Glu Ile Arg Gln Val Gly Asp Phe His Tyr Val Thr Gly  
850 855 860

Met Thr Thr Asp Asn Leu Lys Cys Pro Cys Gln Val Pro Ser Pro Glu  
865 870 875 880

Phe Phe Thr Glu Leu Asp Gly Val Arg Leu His Arg Phe Ala Pro Pro  
885 890 895

Cys Lys Pro Leu Leu Arg Glu Glu Val Ser Phe Arg Val Gly Leu His  
900 905 910

Glu Tyr Pro Val Gly Ser Gln Leu Pro Cys Glu Pro Glu Pro Asp Val  
915 920 925

Ala Val Leu Thr Ser Met Leu Thr Asp Pro Ser His Ile Thr Ala Glu  
930 935 940

Ala Ala Gly Arg Arg Leu Ala Arg Gly Ser Pro Pro Ser Val Ala Ser  
945 950 955 960

Ser Ser Ala Ser Gln Leu Ser Ala Pro Ser Leu Lys Ala Thr Cys Thr  
965 970 975

Ala Asn His Asp Ser Pro Asp Ala Glu Leu Ile Glu Ala Asn Leu Leu  
980 985 990

Trp Arg Gln Glu Met Gly Gly Asn Ile Thr Arg Val Glu Ser Glu Asn  
995 1000 1005

Lys Val Val Ile Leu Asp Ser Phe Asp Pro Leu Val Ala Glu Glu Asp  
1010 1015 1020

Glu Arg Glu Ile Ser Val Pro Ala Glu Ile Leu Arg Lys Ser Arg Arg  
1025 1030 1035 1040

Phe Ala Gln Ala Leu Pro Val Trp Ala Arg Pro Asp Tyr Asn Pro Pro  
1045 1050 1055

Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr Glu Pro Pro Val Val His  
1060 1065 1070

Gly Cys Pro Leu Pro Pro Lys Ser Pro Pro Val Pro Pro Pro Arg  
1075 1080 1085

Lys Lys Arg Thr Val Val Leu Thr Glu Ser Thr Leu Ser Thr Ala Leu  
1090 1095 1100

Ala Glu Leu Ala Thr Arg Ser Phe Gly Ser Ser Ser Thr Ser Gly Ile  
1095 1110 1115 1120

Thr Gly Asp Asn Thr Thr Ser Ser Glu Pro Ala Pro Ser Gly Cys  
1125 1130 1135

Pro Pro Asp Ser Asp Ala Glu Ser Tyr Ser Ser Met Pro Pro Leu Glu  
1140 1145 1150

Gly Glu Pro Gly Asp Pro Asp Leu Ser Asp Gly Ser Trp Ser Thr Val  
1155 1160 1165

Ser Ser Glu Ala Asn Ala Glu Asp Val Val Cys Cys Ser Met Ser Tyr  
1170 1175 1180

Ser Trp Thr Gly Ala Leu Val Thr Pro Cys Ala Ala Glu Glu Gln Lys  
1185 1190 1195 1200

Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu Arg His His Asn Leu  
1205 1210 1215

Val Tyr Ser Thr Thr Ser Arg Ser Ala Cys Gln Arg Gln Lys Lys Val  
 1220 1225 1230

Thr Phe Asp Arg Leu Gln Val Leu Asp Ser His Tyr Gln Asp Val Leu  
 1235 1240 1245

Lys Glu Val Lys Ala Ala Ser Lys Val Lys Ala Asn Leu Leu Ser  
 1250 1255 1260

Val Glu Glu Ala Cys Ser Leu Thr Pro Pro His Ser Ala Lys Ser Lys  
 265 1270 1275 1280

Phe Gly Tyr Gly Ala Lys Asp Val Arg Cys His Ala Arg Lys Ala Val  
 1285 1290 1295

Thr His Ile Asn Ser Val Trp Lys Asp Leu Leu Glu Asp Asn Val Thr  
 1300 1305 1310

Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu Val Phe Cys Val Gln  
 1315 1320 1325

Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu Ile Val Phe Pro Asp  
 1330 1335 1340

Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu Tyr Asp Val Val Thr  
 345 1350 1355 1360

Lys Leu Pro Leu Ala Val Met Gly Ser Ser Tyr Gly Phe Gln Tyr Ser  
 1365 1370 1375

Pro Gly Gln Arg Val Glu Phe Leu Val Gln Ala Trp Lys Ser Lys Lys  
 1380 1385 1390

Thr Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe Asp Ser Thr Val  
 1395 1400 1405

Thr Glu Ser Asp Ile Arg Thr Glu Glu Ala Ile Tyr Gln Cys Cys Asp  
 1410 1415 1420

Leu Asp Pro Gln Ala Arg Val Ala Ile Lys Ser Leu Thr Glu Arg Leu  
 425 1430 1435 1440

Tyr Val Gly Gly Pro Leu Thr Asn Ser Arg Gly Glu Asn Cys Gly Tyr  
 1445 1450 1455

Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr Ser Cys Gly Asn Thr  
 1460 1465 1470

Leu Thr Cys Tyr Ile Lys Ala Arg Ala Ala Cys Arg Ala Ala Gly Leu  
 1475 1480 1485

Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp Leu Val Val Ile Cys  
 1490 1495 1500

Glu Ser Ala Gly Val Gln Glu Asp Ala Ala Ser Leu Arg Ala Phe Thr  
 505 1510 1515 1520

Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp Pro Pro Gln Pro  
1525 1530 1535

Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser Asn Val Ser Val  
1540 1545 1550

Ala His Asp Gly Ala Gly Lys Arg Val Tyr Tyr Leu Thr Arg Asp Pro  
1555 1560 1565

Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr Ala Arg His Thr Pro  
1570 1575 1580

Val Asn Ser Trp Leu Gly Asn Ile Ile Met Phe Ala Pro Thr Leu Trp  
585 1590 1595 1600

Ala Arg Met Ile Leu Met Thr His Phe Phe Ser Val Leu Ile Ala Arg  
1605 1610 1615

Asp Gln Leu Glu Gln Ala Leu Asp Cys Glu Ile Tyr Gly Ala Cys Tyr  
1620 1625 1630

Ser Ile Glu Pro Leu Asp Leu Pro Pro Ile Ile Gln Arg Leu His Gly  
1635 1640 1645

Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro Gly Glu Ile Asn Arg  
1650 1655 1660

Val Ala Ala Cys Leu Arg Lys Leu Gly Val Pro Pro Leu Arg Ala Trp  
665 1670 1675 1680

Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu Leu Ala Arg Gly Gly  
1685 1690 1695

Arg Ala Ala Ile Cys Gly Lys Tyr Leu Phe Asn Trp Ala Val Arg Thr  
1700 1705 1710

Lys Leu Lys Leu Thr Pro Ile Ala Ala Ala Gly Gln Leu Asp Leu Ser  
1715 1720 1725

Gly Trp Phe Thr Ala Gly Tyr Ser Gly Gly Asp Ile Tyr His Ser Val  
1730 1735 1740

Ser His Ala Arg Pro Arg Trp Ile Trp Phe Cys Leu Leu Leu Ala  
745 1750 1755 1760

Ala Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg Met Ser Thr Asn Pro  
1765 1770 1775

Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Gln Asp  
1780 1785 1790

Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Gly Val Tyr Leu Leu  
1795 1800 1805

Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala Thr Arg Lys Thr Ser  
1810 1815 1820

Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro Ile Pro Lys Ala Arg  
825 1830 1835 1840

Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly Tyr Pro Trp Pro Leu  
1845 1850 1855

Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp Leu Leu Ser Pro Arg  
1860 1865 1870

Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro Arg Arg Arg Ser Arg  
1875 1880 1885

Asn Leu Gly Lys  
1890

<210> 14  
<211> 20316  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
pd.delta.NS3NS5.pj.core173

<220>  
<221> CDS  
<222> (12679) .. (18510)

<400> 14  
atcgatccta ccccttgcgc taaagaagta tatgtgccta ctaacgcttg tctttgtctc 60  
tgtcaactaaa cactggatta ttactcccag atacttattt tggactaatt taaatgattt 120  
cgatatcaacg ttcttaatat cgctgaatct tccacaattt atgaaagtag cttaggaagag 180  
gaattggat aaagttttt tttttgtaaa tctcgaagta tactcaaacg aatttagtat 240  
tttctcagtg atctcccaga tgctttcacc ctcaacttata agtgctttaa gcatttttt 300  
actgtggcta tttcccttat ctgcttcctt ccatgattcg aactgttaatt gcaaactact 360  
tacaatatca gtgatatacg attgatgttc ttgtccatag taaggaataa ttgtaaattc 420  
ccaaggcagga atcaatttct ttaatgaggc ttccagaatt gttgctttt gcgtcttgta 480  
tttaaactgg agtgatttat tgacaatatac gaaactcagc gaattgctta tgatagtatt 540  
atagctcatg aatgtggctc tcttgattgc tgttccgtta tgtgtaatca tccaaacataa 600  
ataggtagt tcagcagcac ataatgctat tttctcacct gaaggtctt caaaccttc 660  
cacaactga cgaacaagca ctttaggtgg tgttttacat aatataatcaa attgtggcat 720  
gcttagcgcc gatcttgtgt gcaattgata tctagttca actactctat ttatcttgta 780  
tcttgcagta ttcaaaacacg ctaactcgaa aaactaactt taattgtcct gtttgcctcg 840

cgttcttcg aaaaatgcac cggccgcga ttatttgtac tgcgaaaata attggtaactg 900  
cggtatctc atttcatatt taaaaatgc acctttgctg ctttccta attttagac 960  
ggccgcagg ttcgtttgc ggtactatct tgtataaaa agttgtttg acaagtatgc 1020  
tgcacagatt ttataatgta ataagcaaga atacattatc aaacgaacaa tactggtaaa 1080  
agaaaaccaa aatggacgac attgaaacag ccaagaatct gacggtaaaa gcacgtacag 1140  
cttatacggt ctggatgta tgctggctgt ttattgaaat gattgctcct gatgtagata 1200  
ttgatataaga gagtaaacgt aagtctgatg agctactctt tccaggatc gtcataaggc 1260  
ccatggaatc tctcacaacc ggtaggccgt atggcttga ttctagcgca gaagattcca 1320  
gcgtatctc tgactccagt gctgaggtaa tttgcctgc tgcaagatg gttaaggaaa 1380  
ggttgattc gattgaaat ggtatgctc cttcacaaga agcaagtcag gctgccatag 1440  
atttgcgt acagaataac aagctgttag acaatagaaa caactatac aaatctattg 1500  
ctataataat aggaagattt cccgagaaag acaagaagag agctaccgaa atgctcatga 1560  
gaaaaatgga ttgtacacag ttattagtcc caccagctcc aacggaagaa gatgttatga 1620  
agctcgtaag cgctgttacc caattgctt ctttagttcc accagatcgt caagctgctt 1680  
taataggtga tttattcatc ccggaatctc taaaggatc attcaatagt ttcaatgaac 1740  
tggccgcaga gaatcgaaa cagaaaaaa agatgagtt ggaaggaagg actgaagtga 1800  
accatgctaa tacaaatgaa gaagttccct ccaggcgaac aagaagtaga gacacaaatg 1860  
caagaggagc atataaatta caaaacacca tcactgaggg ccctaaagcg gttcccacga 1920  
aaaaaaggag agtagcaacg aggtaaggg gcagaaaatc acgtaataact tctagggtat 1980  
gatccatat caaaggaaat gatagcatgg aaggatgaga ctaatccat tgaggagtgg 2040  
cagcatatag aacagctaaa ggtagtgcgtaagca tacgataccc cgcatggaat 2100  
gggataatat cacaggaggt actagactac ctttcatcct acataaatacg acgcatataa 2160  
gtacgcattt aagcataaac acgcactatg ccgttcttct catgtatata tatatacagg 2220  
caacacgcag atatagggtgc gacgtgaaca gtgagctgta tgtgcgcagc tcgcgttgca 2280  
tttcggaaag cgctcgaaaa cggaaacgct ttgaagttcc tattccgaag ttcctattct 2340  
ctagaaaatc taggaacttc agagcgctt tgaaaaccaa aagcgctctg aagacgcact 2400  
ttcaaaaaac caaaaacgcac ccggactgta acgagctact aaaatattgc gaataccgct 2460  
tccacaaaca ttgctcaaaa gtatctctt gctatataatc tctgtgctat atccctat 2520  
aacctaccca tccaccccttc gtccttgaa cttgcataacta aactcgacct ctacatcaac 2580

BIOINFORMATICS

aggcttccaa tgctctcaa atttactgt caagtagacc catacggtg taatatgctg 2640  
ctcttcataa tgtaagctta tctttatcga atcgtgtgaa aaactactac cgcgataaac 2700  
ctttacgggtt ccctgagatt gaatttagttc cttagtata tgataacaaga cactttgaa 2760  
cttgcgtacga cgaatttgaa gttcgccat cctctggcta tttccaatta tcctgtcggc 2820  
tattatctcc gcctcagttt gatctccgc ttcaagactgc cattttcac ataatgaatc 2880  
tatttcaccc cacaatcctt catccgcctc cgcatctgt tccgttaaac tattgacttc 2940  
atgttgtaca ttgttagtt cacgagaagg gtcctttca ggcggtagct cctgatctcc 3000  
tatatgacct ttatcctgtt ctcttccac aaacttagaa atgtattcat gaattatgga 3060  
gcacctaata acattttca aggccggagaa gttggccca gatgccaaat atgcttgaca 3120  
tgaaaacgtg agaatgaatt tagtattatt gtgatattct gaggcaattt tattataatc 3180  
tcgaagataa gagaagaatg cagtgacctt tgtattgaca aatggagatt ccatgtatct 3240  
aaaaaaatacg cctttaggcc ttctgatacc ctccccctg cggttagcg tgcctttac 3300  
attaatatct aaaccctctc cgatggtggc cttaactga ctaataatg caaccgatat 3360  
aaactgtgat aattctgggt gatttatgat tcgatcgaca attgtattgt acactagtgc 3420  
aggatcaggg caatccagtt cttttcaat taccgggtgt tcgtctgtat tcagtacatg 3480  
tccaacaaat gcaaatgcta acgttttta tttcttataa ttgtcaggaa ctggaaaagt 3540  
cccccttgcgtc gtctcgatta cacacctact ttcatcgatc accatagggtt ggaagtgctg 3600  
cataatacat tgcttaatac aagcaagcag tctctcgcca ttcatatttc agttatttc 3660  
cattacagct gatgtcattt gatatcagcg ctgtaaaaat ctatctgtt cagaaggaaa 3720  
tcgcggtttt tataaacaaa actttcgta cgaatcgag caatcccc agctgcgtat 3780  
ttggaaattt gggaaaaagt agagcaacgc gagttgcatt ttttacacca taatgcgtg 3840  
ttaacttcga gaagggatta aggctaattt cactagtatg tttcaaaaac ctcaatctgt 3900  
ccattgaatg ctttataaaa cagctataga ttgcataagaa gagttagcta ctcaatgctt 3960  
tttgtcaaag cttactgtatg atgatgtgtc tacttcagg cgggtctgtt gtaaggagaa 4020  
tgacattata aagctggcac ttgtttttcc acggactata gactatacta gtatactccg 4080  
tctactgtac gatacacttc cgctcaggc cttgccttt aacgaggcct taccacttt 4140  
ttgttactctt attgtatccag ctcagcaaaag gcagtgtgat ctaagattct atcttcgcga 4200  
tgtatggaaa ctatgtatgc cgagaaagag actagaaatg caaaaggcac ttctacaatg 4260  
gctgccatca ttattatccg atgtgacgct gcatttttt tttttttttt tttttttttt 4320

ttttttttt tttttttttt ttttttgta caaatatcat aaaaaaaagag aatctttta 4380  
agcaaggatt ttcttaactt ctgcggcgac agcatcacccg acttcggtgg tactgttgg 4440  
accacctaata tcaccagtgc tgataacctgc atccaaaacc tttttactg catcttaat 4500  
ggcttaccc tcttcaggca agttcaatga caattcaac atcattgcag cagacaagat 4560  
agtggcgata gggttgacct tattcttgg caaatctgga gcggaaaccat ggcattgttc 4620  
gtacaaacca aatgcggtgt tcttgcgtgg caaagaggcc aaggacgcag atggcaacaa 4680  
acccaaggag cctggataa cgaggcttc atcgagatg atatcaccaa acatgttgct 4740  
ggtgattata ataccattta ggtgggttgg gttcttaact aggatcatgg cggcagaatc 4800  
aatcaattga tggtaactt tcaatgttagg gaattcggtt ttgatgggtt cctccacagt 4860  
tttctccat aatctgaag aggccaaaac attagctta tccaaggacc aaataggcaa 4920  
tggtggtca tggtaaggccatgaaagc ggccattttt gtgattttt gcacttctgg 4980  
aacgggttat tggtaactat cccagcgac accatcacca tcgttttctt ttcttttacc 5040  
aaagtaaata cctccacta attctctaac aacaacgaag tcgtacctt tagcaaattg 5100  
tggcttgatt ggagataagt ctaaaagaga gtcggatgca aagttacatg gtcttaagtt 5160  
ggcgtacaat tggtaactt tacggatttt tagtaaacct tggcaggc taacactacc 5220  
ggtagccccat ttaggaccac ccacagcacc taacaaaacg gcatcaggct tcttgaggc 5280  
ttccagcgcc tcatctggaa gtggAACACC tggatcgatcg atagcagcac caccaattaa 5340  
atgattttcg aaatcgaact tggatcgatcg acgaacatca gaaatagctt taagaacctt 5400  
aatggcttcg gctgtgattt cttggatcgatcg acgtttttttt ggcaaaacga cgatcttctt 5460  
aggggcagac attacaatgg tatatccttg aaaaaaaaaaaaaaaa 5520  
aaaaaaaaaaa atgcagcttc tcaatgtat tggatcgatcg tttggaggaga tacagctaa 5580  
tatccgacaa actgttttac agatttacga tcgtacttgc taccatcat tggatcgatcg 5640  
acatccgaac ctggagttt tccctgaaac agatgtata tttggatcgatcg tataataata 5700  
tatagtcgtatcg cgcttacgg aagacaatgt atgtatcg tggatcgatcg aaactattgc 5760  
atctattgca taggtatct tggatcgatcg atccccgggtt cattttctgc gtttccatct 5820  
tgcacttcaa tagcatatct tggatcgatcg agcatctgtg cttcattttt tagaaca 5880  
atgcacgcgc agagcgctaa tttttcaac aaaaatcgatcg agctgcattt ttacagaaca 5940  
gaaatgcac gcaaaagcgc tttttacca acgaagaatc tggatcgatcg tttttttttt 6000  
caaaaatgcac acgcgagac gcttacattttt caaaca 6060

gaacagaaaat gcaacgcgag agcgctattt taccaacaaa gaatctatac ttctttttg 6120  
ttctacaaaa atgcatccc agagcgctat ttttctaaca aagcatctta gattacttt 6180  
tttctccctt gtgcgctcta taatgcagtc tcttgataac ttttgcact gtaggtccgt 6240  
taaggtaga agaaggctac tttgggtgtct attttctctt ccataaaaaaa agcctgaactc 6300  
cacttccccgc gtttactgat tactagcgaa gctgcgggtg catttttca agataaaggc 6360  
atccccgatt atattctata ccgatgtgga ttgcgcatac tttgtgaaca gaaagtgata 6420  
gcgttcatgtatc ttcttcattt gtcagaaaat tatgaacggt ttcttctatt ttgtcttat 6480  
atactacgtatc taggaaatgt ttacatttc gtattgttt cgattcactc tatgaatagt 6540  
tcttactaca attttttgt cttaaagagta atactagaga taaacataaa aaatgttagag 6600  
gtcgagttt gatgcaagtt caaggagcga aagggtggatg ggttaggttat atagggatat 6660  
agcacagaga tatatacgaa agagatactt ttgagcaatg tttgttggaaag cggtattcgc 6720  
aatatttttag tagctcgatc cagtcgggtg cgtttttgtt ttgttggaaag tgcttcttca 6780  
gagcgctttt ggtttcaaa agcgctctga agttcctata ctttcttagag aataggaact 6840  
tcggaatagg aacttcaaag cgtttccgaa aacgagcgt tccgaaaatg caacgcgagc 6900  
tgcgcacata cagctcactg ttcacgtcgc acctatatct gcgtgttgcc tgtatataata 6960  
tatacatgag aagaacggca tagtgcgtgt ttatgcttaa atgcgtactt atatgcgtct 7020  
atttatgttag gatgaaaggt agtctagtagtact ctoctgtatc attatcccat tccatgcggg 7080  
gtatcgatgtatc cttcccttcag cactaccctt tagctgttct atatgctgcc actcctcaat 7140  
tggattagtc tcatccctca atgctatcat ttcctttatc attggatcat atgcatacgta 7200  
ccgagaaact agtgcgaagt agtgatcagg tattgctgtt atctgatgag tatacggtt 7260  
cctggccacg gcagaagcac gcttatacgct ccaatttccc acaacattag tcaactccgt 7320  
taggcccttc attgaaagaa atgaggtcat caaatgtctt ccaatgttagt attttgggcc 7380  
attttttata gcaaagattt aataaggcgc atttttcttc aaagctttat tgcgtatct 7440  
gactaagtta tcttttaata attggatttc ctgtttattt cttgaagaat tgccggccct 7500  
atttactcgatc tttaggactgtt cttcagaatt cctcaaaaat tcatccaaat atacaagtgg 7560  
atcgatgata agctgtcaaa catgagaatt cttgaagacg aaagggcctc gtgatacgcc 7620  
tattttata ggttaatgtc atgataataa tggtttctta gacgtcaggt ggcacttttc 7680  
ggggaaatgt gcgccgaaacc cctatttgcattttctta aatacattca aatatgtatc 7740  
cgctcatgag acaataaaccctgataatgc ttcaataataa ttgaaaaagg aagagtatga 7800

gtattcaaca tttccgtgtc gcccttattc ccttttgc ggcatttgc cttcctgttc 7860  
ttgctcaccc agaaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag 7920  
tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagttt cgccccgaag 7980  
aacgtttcc aatgatgagc acttttaaag ttctgctatg tggcgcgta ttatcccgtg 8040  
ttgacgcccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg 8100  
agtactcacc agtcacagaa aagcatctt cggatggcat gacagtaaga gaattatgca 8160  
gtgctgccat aaccatgagt gataaacactg cggccaactt acttctgaca acgatcgag 8220  
gaccgaagga gctaaccgct ttttgacaca acatggggga tcatgtaact cgccttgatc 8280  
gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg 8340  
cagcaatggc aacaacgttg cgcaaactat taactggcga actacttact ctagcttccc 8400  
ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg 8460  
cccttccggc tggctggttt attgctgata aatctggagc cggtagcgt gggtctcgcg 8520  
gtatcattgc agcaactgggg ccagatggta agccctcccg tatcgtagtt atctacacga 8580  
cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac 8640  
tgattaagca ttggtaactg tcagaccaag tttactcata tatacttttag attgatttaa 8700  
aacttcattt ttaattnaa aggatctagg tgaagatctt tttgataat ctcatgacca 8760  
aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag 8820  
gatcttctt agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaaccac 8880  
cgctaccagc ggtgggttgt ttgcccgtc aagagctacc aactctttt ccgaaggtaa 8940  
ctggcttcag cagagcgcag ataccaaata ctgtccttct agttagccg tagttaggcc 9000  
accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaatc ctgttaccag 9060  
tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac 9120  
cggtataaggc gcagcggcgtg ggctgaacgg ggggttcgtg cacacagccc agcttggagc 9180  
gaacgaccta caccgaactg agataccac agcgtgagct atgagaaagc gccacgcttc 9240  
ccgaaggagaa aaaggcggac aggtatccgg taagcggcag ggtcggaaca ggagagcgca 9300  
cgagggagct tccaggggaa aacgcctgtt atctttatag tcctgtcggtt tttcgccacc 9360  
tctgacttga gcgtcgatatt ttgtgtatgtc cgtcaggggg gctggaccta tggaaaaaacg 9420  
ccagcaacgc ggcctttta cgttccctgg cctttgctg gcctttgct cacatgttct 9480  
ttcctgcgtt atccctgtat cgtgtggata accgttattac cgccttgag tgagctgata 9540

ccgctcgccg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa gcggaagac 9600  
gcctgatgcg gtattttctc cttacgcac ttgtcggtat ttcacaccgc atatggtca 9660  
ctctcagtagtac aatctgctct gatgccgcat agttaagcca gtatacactc cgctatcgct 9720  
acgtgactgg gtcatggctg cgccccgaca cccgccaaca cccgctgacg cgcctgacg 9780  
ggcttgctg ctcccccgt catcaccgaa acgcgcgagg cagctgcggt aaagctcatc 9840  
gtgtcagagg ttttaccgt catcaccgaa acgcgcgagg cagctgcggt aaagctcatc 9900  
acggtggtcg tgaagcgatt cacagatgtc tgccgttca tccgcgtcca gtcgtttag 9960  
tttctccaga agcgtaatg tctggcttct gataaagcgg gccatgttaa gggcggttt 10020  
ttcctgtttg gtcactgatg cctccgtgt aaaaaaaaaaaaaaaaatgatgatgat 10080  
accgatgaaa cgagagagga tgctcacgt acgggttact gatgtgaac atgcccggtt 10140  
actggAACGT tgtgagggt aacaactggc ggtatggatg cggcgggacc agagaaaaat 10200  
cactcagggt caatgccagc gcttcgttaa tacagatgtc ggttccac agggtagcca 10260  
gcagcatcct gcgatgcaga tccggaacat aatggtcag ggcgtgact tccgcgtttc 10320  
cagactttac gaaacacgga aaccgaagac cattcatgtt gttgtcagg tcgcagacgt 10380  
tttgcagcag cagtcgcttc acgttcgctc gcgtatcggt gattcattct gctaaccagt 10440  
aaggcaaccc cgccagccta gccgggtctt caacgacagg agcacgtca tgccacccg 10500  
tggccaggac ccaacgctgc ccgagatgcg ccgcgtgcgg ctgctggaga tggcggacgc 10560  
gatggatatg ttctgccaag gttggtttgcgcattcaca gttctccgca agaattgatt 10620  
ggctccaatt cttggagtgg tgaatccgtt agcgaggtgc cgccggcttc cattcaggc 10680  
gaggtggccc ggctccatgc accgcgacgc aacgcggggaa ggcagacaag gtatagggcg 10740  
gcgcctacaa tccatgcca cccgttccat gtgcgtcccg aggccgcata aatgcgcgt 10800  
acgatcagcg gtccaatgat cgaagttagg ctggtaagag ccgcgagcga tccttgaagc 10860  
tgtccctgat ggctcgatc tacctgcctg gacagcatgg cctgcaacgc gggcatcccg 10920  
atgcgcgg aagcgagaag aatcataatg gggaaaggcca tccagcctcg cgtcgac 10980  
gccagcaaga cgtagcccg cgcgtcgccg cccatgcgg cgataatggc ctgcttctcg 11040  
ccgaaacgtt tgggtggcgaa accagtgcacg aaggcttgcg cgagggcggtg caagattccg 11100  
aataccgcaa gcgcacaggcc gatcatgcgcgcgtccacgc gaaagcggtc ctgcggaaa 11160  
atgacccaga gcgcgtccgg cacctgtct acgagttgca tgataaagaa gacagtcata 11220  
agtgcggcga cgatagtcat gccccggcc caccggaaagg agctgactgg gttgaaggct 11280

ctcaaggcga tcggtcgagg atccttcaat atgcgcacat acgctgttat gttcaaggc 11340  
 ccttcgttta agaacgaaag cggtcttcct tttgagggat gttcaagtt gttcaaact 11400  
 atcaaatttg caaatccccca gtctgtatct agagcggtga atcggtgatg cgatggta 11460  
 attaaattga tggtgtcacc attaccaggt ctagatatac caatggcaaa ctgagcacaa 11520  
 caataccagt ccggatcaac tggcaccatc tctcccttag tctcatctaa tttttcttcc 11580  
 ggatgaggtt ccagatatac cgcaacaccc ttattatggt ttccctgagg gaataataga 11640  
 atgtcccatt cgaaatcacc aattctaaac ctggcgaat tgtatttcgg gtttgttaac 11700  
 tcgttccagt caggaatgtt ccacgtgaag ctatcttcca gcaaagtctc cacttcttca 11760  
 tcaaattgtg gagaatactc ccaatgtct tatctatggg acttccggaa aacacagtag 11820  
 cgatacttcc caattcgtct tcagagctca ttgtttgtt gaagagacta atcaaagaat 11880  
 cgtttctca aaaaaattaa tatcttaact gatagttga tcaaaggc 11940  
 ggcaaacaaa cgaaaaatc gtttctcaa tttctgatg ccaagaactc taaccagtct 12000  
 tatctaaaaa ttgcctttag atccgtctct ccgggtacag cctgtgtac tgattaatcc 12060  
 tgccttctca atcaccattc taatgtttta attaagggat tttgtttca ttaacggctt 12120  
 tcgctcataa aaatgttag acgtttgcc cgccggcggg aaaccatcca cttcacgaga 12180  
 ctgatctcct ctgcccgaac accgggcatc tccaacttat aagttggaga aataagagaa 12240  
 tttcagattt agagaatgaa aaaaaaaaaac ctttagttca taggtccatt ctcttagcgc 12300  
 aactacagag aacagggcga caaacaggca aaaaacggc acaacctcaa tggagtgtat 12360  
 caacctgcct ggagtaatg atgacacaag gcaattgacc cacgcatgta tctatctcat 12420  
 tttcttacac cttctattac cttctgtct ctctgatgg aaaaaagctg aaaaaaaaagg 12480  
 ttgaaaccag ttccctgaaa ttattccct acttgactaa taagtatata aagacggtag 12540  
 gtattgattt taattctgtt aatctatttc ttaaacttct taaattctac tttttagtt 12600  
 agtctttttt ttagtttaa aacaccaaga acttagttc gaataaacac acataaaca 12660  
 acaagcttac aaaacaaa atg gct gca tat gca gct cag ggc tat aag gtg 12711  
                   Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val  
                   1              5              10  
  
 cta gta ctc aac ccc tct gtt gct gca aca ctg ggc ttt ggt gct tac 12759  
 Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr  
                   15              20              25  
  
 atg tcc aag gct cat ggg atc gat cct aac atc agg acc ggg gtg aga 12807  
 Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg  
                   30              35              40

aca att acc act ggc agc ccc atc acg tac tcc acc tac ggc aag ttc Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe 45 50 55	12855
ctt gcc gac ggc ggg tgc tcg ggg ggc gct tat gac ata ata att tgt Leu Ala Asp Gly Gly Cys Ser Gly Ala Tyr Asp Ile Ile Ile Cys 60 65 70 75	12903
gac gag tgc cac tcc acg gat gcc aca tcc atc ttg ggc att ggc act Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr 80 85 90	12951
gtc ctt gac caa gca gag act gcg ggg gcg aga ctg gtt gtg ctc gcc Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala 95 100 105	12999
acc gcc acc cct ccg ggc tcc gtc act gtg ccc cat ccc aac atc gag Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn Ile Glu 110 115 120	13047
gag gtt gct ctg tcc acc acc gga gag atc cct ttt tac ggc aag gct Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala 125 130 135	13095
atc ccc ctc gaa gta atc aag ggg ggg aga cat ctc atc ttc tgt cat Ile Pro Leu Glu Val Ile Lys Gly Arg His Leu Ile Phe Cys His 140 145 150 155	13143
tca aag aag aag tgc gac gaa ctc gcc gca aag ctg gtc gca ttg ggc Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Ala Leu Gly 160 165 170	13191
atc aat gcc gtg gcc tac tac cgc ggt ctt gac gtg tcc gtc atc ccg Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile Pro 175 180 185	13239
acc agc ggc gat gtt gtc gtc gtg gca acc gat gcc ctc atg acc ggc Thr Ser Gly Asp Val Val Val Ala Thr Asp Ala Leu Met Thr Gly 190 195 200	13287
tat acc ggc gac ttc gac tcg gtg ata gac tgc aat acg tgt gtc acc Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys Val Thr 205 210 215	13335
cag aca gtc gat ttc agc ctt gac cct acc ttc acc att gag aca atc Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr Ile 220 225 230 235	13383
acg ctc ccc caa gat gct gtc tcc cgc act caa cgt cgg ggc agg act Thr Leu Pro Gln Asp Ala Val Ser Arg Thr Gln Arg Arg Gly Arg Thr 240 245 250	13431
ggc agg ggg aag cca ggc atc tac aga ttt gtg gca ccg ggg gag cgc Gly Arg Gly Lys Pro Gly Ile Tyr Arg Phe Val Ala Pro Gly Glu Arg 255 260 265	13479
ccc tcc ggc atg ttc gac tcg tcc gtc ctc tgt gag tgc tat gac gca	13527

Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp Ala			
270	275	280	
ggc tgt gct tgg tat gag ctc acg ccc gcc gag act aca gtt agg cta			13575
Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Thr Val Arg Leu			
285	290	295	
cga gcg tac atg aac acc ccg ggg ctt ccc gtg tgc cag gac cat ctt			13623
Arg Ala Tyr Met Asn Thr Pro Gly Leu Pro Val Cys Gln Asp His Leu			
300	305	310	315
gaa ttt tgg gag ggc gtc ttt aca ggc ctc act cat ata gat gcc cac			13671
Glu Phe Trp Glu Gly Val Phe Thr Gly Leu Thr His Ile Asp Ala His			
320	325	330	
ttt cta tcc cag aca aag cag agt ggg gag aac ctt cct tac ctg gta			13719
Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Leu Pro Tyr Leu Val			
335	340	345	
gcg tac caa gcc acc gtg tgc gct agg gct caa gcc cct ccc cca tcg			13767
Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser			
350	355	360	
tgg gac cag atg tgg aag tgt ttg att cgd ctc aag ccc acc ctc cat			13815
Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His			
365	370	375	
ggg cca aca ccc ctg cta tac aga ctg ggc gct gtt cag aat gaa atc			13863
Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Ile			
380	385	390	395
acc ctg acg cac cca gtc acc aaa tac atc atg aca tgc atg tcg gcc			13911
Thr Leu Thr His Pro Val Thr Lys Tyr Ile Met Thr Cys Met Ser Ala			
400	405	410	
gac ctg gag gtc acg agc acc tgg gtc ctc gtt ggc ggc gtc ctg			13959
Asp Leu Glu Val Val Thr Ser Thr Trp Val Leu Val Gly Gly Val Leu			
415	420	425	
gct gct ttg gcc gcg tat tgc ctg tca aca ggc tgc gtc gtc ata gtg			14007
Ala Ala Leu Ala Ala Tyr Cys Leu Ser Thr Gly Cys Val Val Ile Val			
430	435	440	
ggc agg gtc gtc ttg tcc ggg aag ccg gca atc ata cct gac agg gaa			14055
Gly Arg Val Val Leu Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu			
445	450	455	
gtc ctc tac cga gag ttc gat gag atg gaa gag tgc tct cag cac tta			14103
Val Leu Tyr Arg Glu Phe Asp Glu Met Glu Glu Cys Ser Gln His Leu			
460	465	470	475
ccg tac atc gag caa ggg atg atg ctc gcc gag cag ttc aag cag aag			14151
Pro Tyr Ile Glu Gln Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys			
480	485	490	
gcc ctc ggc ctc ctg cag acc gcg tcc cgt cag gca gag gtt atc gcc			14199
Ala Leu Gly Leu Leu Gln Thr Ala Ser Arg Gln Ala Glu Val Ile Ala			

495	500	505	
cct gct gtc cag acc aac tgg caa aaa ctc gag acc ttc tgg gcg aag Pro Ala Val Gln Thr Asn Trp Gln Lys Leu Glu Thr Phe Trp Ala Lys			14247
510	515	520	
cat atg tgg aac ttc atc agt ggg ata caa tac ttg gcg ggc ttg tca His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala Gly Leu Ser			14295
525	530	535	
acg ctg cct ggt aac ccc gcc att gct tca ttg atg gct ttt aca gct Thr Leu Pro Gly Asn Pro Ala Ile Ala Ser Leu Met Ala Phe Thr Ala			14343
540	545	550	555
gct gtc acc agc cca cta acc act agc caa acc ctc ctc ttc aac ata Ala Val Thr Ser Pro Leu Thr Ser Gln Thr Leu Leu Phe Asn Ile			14391
560	565	570	
ttg ggg ggg tgg gtg gct gcc cag ctc gcc gcc ccc ggt gcc gct act Leu Gly Gly Trp Val Ala Ala Gln Leu Ala Ala Pro Gly Ala Ala Thr			14439
575	580	585	
gcc ttt gtg ggc gct ggc tta gct ggc gcc gcc atc ggc agt gtt gga Ala Phe Val Gly Ala Gly Leu Ala Gly Ala Ala Ile Gly Ser Val Gly			14487
590	595	600	
ctg ggg aag gtc ctc ata gac atc ctt gca ggg tat ggc gcg ggc gtg Leu Gly Lys Val Leu Ile Asp Ile Leu Ala Gly Tyr Gly Ala Gly Val			14535
605	610	615	
gcg gga gct ctt gtg gca ttc aag atc atg agc ggt gag gtc ccc tcc Ala Gly Ala Leu Val Ala Phe Lys Ile Met Ser Gly Glu Val Pro Ser			14583
620	625	630	635
acg gag gac ctg gtc aat cta ctg ccc gcc atc ctc tcg ccc gga gcc Thr Glu Asp Leu Val Asn Leu Leu Pro Ala Ile Leu Ser Pro Gly Ala			14631
640	645	650	
ctc gta gtc ggc gtg gtc tgt gca gca ata ctg cgc cgg cac gtt ggc Leu Val Val Gly Val Val Cys Ala Ala Ile Leu Arg Arg His Val Gly			14679
655	660	665	
ccg ggc gag ggg gca gtg cag tgg atg aac cgg ctg ata gcc ttc gcc Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile Ala Phe Ala			14727
670	675	680	
tcc cgg ggg aac cat gtt tcc ccc acg cac tac gtg ccg gag agc gat Ser Arg Gly Asn His Val Ser Pro Thr His Tyr Val Pro Glu Ser Asp			14775
685	690	695	
gca gct gcc cgc gtc act gcc ata ctc agc agc ctc act gta acc cag Ala Ala Ala Arg Val Thr Ala Ile Leu Ser Ser Leu Thr Val Thr Gln			14823
700	705	710	715
ctc ctg agg cga ctg cac cag tgg ata agc tcg gag tgt acc act cca Leu Leu Arg Arg Leu His Gln Trp Ile Ser Ser Glu Cys Thr Thr Pro			14871
720	725	730	

tgc tcc ggt tcc tgg cta agg gac atc tgg gac tgg ata tgc gag gtg 14919  
 Cys Ser Gly Ser Trp Leu Arg Asp Ile Trp Asp Trp Ile Cys Glu Val  
 735 740 745

ttg agc gac ttt aag acc tgg cta aaa gct aag ctc atg cca cag ctg 14967  
 Leu Ser Asp Phe Lys Thr Trp Leu Lys Ala Lys Leu Met Pro Gln Leu  
 750 755 760

cct ggg atc ccc ttt gtg tcc tgc cag cgc ggg tat aag ggg gtc tgg 15015  
 Pro Gly Ile Pro Phe Val Ser Cys Gln Arg Gly Tyr Lys Gly Val Trp  
 765 770 775

cga ggg gac ggc atc atg cac act cgc tgc cac tgt gga gct gag atc 15063  
 Arg Gly Asp Gly Ile Met His Thr Arg Cys His Cys Gly Ala Glu Ile  
 780 785 790 795

act gga cat gtc aaa aac ggg acg atg agg atc gtc ggt cct agg acc 15111  
 Thr Gly His Val Lys Asn Gly Thr Met Arg Ile Val Gly Pro Arg Thr  
 800 805 810

tgc agg aac atg tgg agt ggg acc ttc ccc att aat gcc tac acc acg 15159  
 Cys Arg Asn Met Trp Ser Gly Thr Phe Pro Ile Asn Ala Tyr Thr Thr  
 815 820 825

ggc ccc tgt acc ccc ctt cct gcg ccg aac tac acg ttc gcg cta tgg 15207  
 Gly Pro Cys Thr Pro Leu Pro Ala Pro Asn Tyr Thr Phe Ala Leu Trp  
 830 835 840

agg gtg tct gca gag gaa tac gtg gag ata agg cag gtg ggg gac ttc 15255  
 Arg Val Ser Ala Glu Glu Tyr Val Glu Ile Arg Gln Val Gly Asp Phe  
 845 850 855

cac tac gtg acg ggt atg act act gac aat ctt aaa tgc ccg tgc cag 15303  
 His Tyr Val Thr Gly Met Thr Asp Asn Leu Lys Cys Pro Cys Gln  
 860 865 870 875

gtc cca tcg ccc gaa ttt ttc aca gaa ttg gac ggg gtg cgc cta cat 15351  
 Val Pro Ser Pro Glu Phe Phe Thr Glu Leu Asp Gly Val Arg Leu His  
 880 885 890

agg ttt gcg ccc ccc tgc aag ccc ttg ctg cgg gag gag gta tca ttc 15399  
 Arg Phe Ala Pro Pro Cys Lys Pro Leu Leu Arg Glu Glu Val Ser Phe  
 895 900 905

aga gta gga ctc cac gaa tac ccg gta ggg tcg caa tta cct tgc gag 15447  
 Arg Val Gly Leu His Glu Tyr Pro Val Gly Ser Gln Leu Pro Cys Glu  
 910 915 920

ccc gaa ccg gac gtg gcc gtg ttg acg tcc atg ctc act gat ccc tcc 15495  
 Pro Glu Pro Asp Val Ala Val Leu Thr Ser Met Leu Thr Asp Pro Ser  
 925 930 935

cat ata aca gca qag qcgc ggc ggg cga agg ttg gcg agg gga tca ccc 15543  
 His Ile Thr Ala Glu Ala Ala Gly Arg Arg Leu Ala Arg Gly Ser Pro  
 940 945 950 955

ccc tct gtg gcc agc tcc tcg gct agc cag cta tcc gct cca tct ctc 15591

Pro Ser Val Ala Ser Ser Ser Ala Ser Gln Leu Ser Ala Pro Ser Leu		
960	965	970
aag gca act tgc acc gct aac cat gac tcc cct gat gct gag ctc ata		15639
Lys Ala Thr Cys Thr Ala Asn His Asp Ser Pro Asp Ala Glu Leu Ile		
975	980	985
gag gcc aac ctc cta tgg agg cag gag atg ggc ggc aac atc acc agg		15687
Glu Ala Asn Leu Leu Trp Arg Gln Glu Met Gly Gly Asn Ile Thr Arg		
990	995	1000
gtt gag tca gaa aac aaa gtg gtg att ctg gac tcc ttc gat ccg ctt		15735
Val Glu Ser Glu Asn Lys Val Val Ile Leu Asp Ser Phe Asp Pro Leu		
1005	1010	1015
gtg gcg gag gag gac gag cg <sup>g</sup> gag atc tcc gta ccc gca gaa atc ctg		15783
Val Ala Glu Glu Asp Glu Arg Glu Ile Ser Val Pro Ala Glu Ile Leu		
1020	1025	1030
1035		
cgg aag tct cgg aga ttc gcc cag gcc ctg ccc gtt tgg gcg cgg c <sup>g</sup> c <sup>g</sup>		15831
Arg Lys Ser Arg Arg Phe Ala Gln Ala Leu Pro Val Trp Ala Arg Pro		
1040	1045	1050
gac tat aac ccc ccg cta gtg gag acg tgg aaa aag ccc gac tac gaa		15879
Asp Tyr Asn Pro Pro Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr Glu		
1055	1060	1065
cca cct gtg gtc cat ggc tgc ccg ctt cca cct cca aag tcc cct cct		15927
Pro Pro Val Val His Gly Cys Pro Leu Pro Pro Pro Lys Ser Pro Pro		
1070	1075	1080
gtg cct ccg cct ccg aag aag cgg acg gtg gtc ctc act gaa tca acc		15975
Val Pro Pro Pro Arg Lys Lys Arg Thr Val Val Leu Thr Glu Ser Thr		
1085	1090	1095
cta tct act gcc ttg gcc gag ctc gcc acc aga agc ttt ggc agc tcc		16023
Leu Ser Thr Ala Leu Ala Glu Leu Ala Thr Arg Ser Phe Gly Ser Ser		
1100	1105	1110
1115		
tca act tcc ggc att acg ggc gac aat acg aca aca tcc tct gag ccc		16071
Ser Thr Ser Gly Ile Thr Gly Asp Asn Thr Thr Ser Ser Glu Pro		
1120	1125	1130
gcc cct tct ggc tgc ccc ccc gac tcc gac gct gag tcc tat tcc tcc		16119
Ala Pro Ser Gly Cys Pro Pro Asp Ser Asp Ala Glu Ser Tyr Ser Ser		
1135	1140	1145
atg ccc ccc ctg gag ggg gag cct ggg gat ccg gat ctt agc gac ggg		16167
Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu Ser Asp Gly		
1150	1155	1160
tca tgg tca acg gtc agt agt gag gcc aac gcg gag gat gtc gtg tgc		16215
Ser Trp Ser Thr Val Ser Ser Glu Ala Asn Ala Glu Asp Val Val Cys		
1165	1170	1175
tgc tca atg tct tac tct tgg aca ggc gca ctc gtc acc ccg tgc gcc		16263
Cys Ser Met Ser Tyr Ser Trp Thr Gly Ala Leu Val Thr Pro Cys Ala		

1180	1185	1190	1195	
gcg gaa gaa cag aaa ctg ccc atc aat gca cta agc aac tcg ttg cta Ala Glu Glu Gln Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu				16311
1200	1205	1210		
cgt cac cac aat ttg gtg tat tcc acc acc tca cgc agt gct tgc caa Arg His His Asn Leu Val Tyr Ser Thr Ser Arg Ser Ala Cys Gln				16359
1215	1220	1225		
agg cag aag aaa gtc aca ttt gac aga ctg caa gtt ctg gac agc cat Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Ser His				16407
1230	1235	1240		
tac cag gac gta ctc aag gag gtt aaa gca gcg gtc tca aaa gtg aag Tyr Gln Asp Val Leu Lys Glu Val Lys Ala Ala Ser Lys Val Lys				16455
1245	1250	1255		
gct aac ttg cta tcc gta gag gaa gct tgc agc ctg acg ccc cca cac Ala Asn Leu Leu Ser Val Glu Glu Ala Cys Ser Leu Thr Pro Pro His				16503
1260	1265	1270	1275	
tca gcc aaa tcc aag ttt ggt tat ggg gca aaa gac gtc cgt tgc cat Ser Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Cys His				16551
1280	1285	1290		
gcc aga aag gcc gta acc cac atc aac tcc gtc tgg aaa gac ctt ctg Ala Arg Lys Ala Val Thr His Ile Asn Ser Val Trp Lys Asp Leu Leu				16599
1295	1300	1305		
gaa gac aat gta aca cca ata gac act acc atc atg gct aag aac gag Glu Asp Asn Val Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu				16647
1310	1315	1320		
gtt ttc tgc gtt cag cct gag aag ggg ggt cgt aag cca gct cgt ctc Val Phe Cys Val Gln Pro Glu Lys Gly Arg Lys Pro Ala Arg Leu				16695
1325	1330	1335		
atc gtg ttc ccc gat ctg ggc gtg cgc gtg tgc gaa aag atg gct ttg Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu				16743
1340	1345	1350	1355	
tac gac gtg gtt aca aag ctc ccc ttg gcc gtg atg gga agc tcc tac Tyr Asp Val Val Thr Lys Leu Pro Leu Ala Val Met Gly Ser Ser Tyr				16791
1360	1365	1370		
gga ttc caa tac tca cca gga cag cgg gtt gaa ttc ctc gtg caa gcg Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Gln Ala				16839
1375	1380	1385		
tgg aag tcc aag aaa acc cca atg ggg ttc tcg tat gat acc cgc tgc Trp Lys Ser Lys Lys Thr Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys				16887
1390	1395	1400		
ttt gac tcc aca gtc act gag agc gac atc cgt acg gag gag gca atc Phe Asp Ser Thr Val Thr Glu Ser Asp Ile Arg Thr Glu Glu Ala Ile				16935
1405	1410	1415		

tac caa tgt tgt gac ctc gac ccc caa gcc cgc gtg gcc atc aag tcc		16983
Tyr Gln Cys Cys Asp Leu Asp Pro Gln Ala Arg Val Ala Ile Lys Ser		
1420 1425 1430 1435		
ctc acc gag agg ctt tat gtt ggg ggc cct ctt acc aat tca agg ggg		17031
Leu Thr Glu Arg Leu Tyr Val Gly Gly Pro Leu Thr Asn Ser Arg Gly		
1440 1445 1450		
gag aac tgc ggc tat cgc agg tgc cgc gcg agc ggc gta ctg aca act		17079
Glu Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr		
1455 1460 1465		
agc tgt ggt aac acc ctc act tgc tac atc aag gcc cggt gca gcc tgt		17127
Ser Cys Gly Asn Thr Leu Thr Cys Tyr Ile Lys Ala Arg Ala Ala Cys		
1470 1475 1480		
cga gcc gca ggg ctc cag gac tgc acc atg ctc gtgt ggc gac gac		17175
Arg Ala Ala Gly Leu Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp		
1485 1490 1495		
tta gtc gtt atc tgt gaa agc gcg ggg gtc cag gag gac gcg gcg agc		17223
Leu Val Val Ile Cys Glu Ser Ala Gly Val Gln Glu Asp Ala Ala Ser		
1500 1505 1510 1515		
ctg aga gcc ttc acg gag gct atg acc agg tac tcc gcc ccc cct ggg		17271
Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly		
1520 1525 1530		
gac ccc cca caa cca gaa tac gac ttg gag ctc ata aca tca tgc tcc		17319
Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser		
1535 1540 1545		
tcc aac gtg tca gtc gcc cac gac ggc gct gga aag agg gtc tac tac		17367
Ser Asn Val Ser Val Ala His Asp Gly Ala Gly Lys Arg Val Tyr Tyr		
1550 1555 1560		
ctc acc cgt gac cct aca acc ccc ctc gcg aga gct gcg tgg gag aca		17415
Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr		
1565 1570 1575		
gca aga cac act cca gtc aat tcc tgg cta ggc aac ata atc atg ttt		17463
Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Phe		
1580 1585 1590 1595		
gcc ccc aca ctg tgg gcg agg atg ata ctg atg acc cat ttc ttt agc		17511
Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser		
1600 1605 1610		
gtc ctt ata gcc agg gac cag ctt gaa cag gcc ctc gat tgc gag atc		17559
Val Leu Ile Ala Arg Asp Gln Leu Glu Gln Ala Leu Asp Cys Glu Ile		
1615 1620 1625		
tac ggg gcc tgc tac tcc ata gaa cca ctg gat cta cct cca atc att		17607
Tyr Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro Pro Ile Ile		
1630 1635 1640		
caa aga ctc cat ggc ctc agc gca ttt tca ctc cac agt tac tct cca		17655

Gln Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro  
 1645 1650 1655

ggt gaa atc aat agg gtg gcc gca tgc ctc aga aaa ctt ggg gta ccg 17703  
 Gly Glu Ile Asn Arg Val Ala Ala Cys Leu Arg Lys Leu Gly Val Pro  
 1660 1665 1670 1675

ccc ttg cga gct tgg aga cac cg<sup>g</sup> gcc cg<sup>g</sup> agc gtc cg<sup>c</sup> gct agg ctt 17751  
 Pro Leu Arg Ala Trp Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu  
 1680 1685 1690

ctg gcc aga gga ggc agg gct gcc ata tgt ggc aag tac ctc ttc aac 17799  
 Leu Ala Arg Gly Gly Arg Ala Ala Ile Cys Gly Lys Tyr Leu Phe Asn  
 1695 1700 1705

tgg gca gta aga aca aag ctc aaa ctc act cca ata gc<sup>g</sup> gcc gct ggc 17847  
 Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Ala Ala Ala Gly  
 1710 1715 1720

cag ctg gac ttg tcc ggc tgg ttc acg gct ggc tac agc ggg gga gac 17895  
 Gln Leu Asp Leu Ser Gly Trp Phe Thr Ala Gly Tyr Ser Gly Gly Asp  
 1725 1730 1735

att tat cac agc gtg tct cat gcc cg<sup>g</sup> ccc cg<sup>c</sup> tgg atc tgg ttt tgc 17943  
 Ile Tyr His Ser Val Ser His Ala Arg Pro Arg Trp Ile Trp Phe Cys  
 1740 1745 1750 1755

cta ctc ctg ctt gct gca ggg gta ggc atc tac ctc ctc ccc aac cga 17991  
 Leu Leu Leu Ala Ala Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg  
 1760 1765 1770

atg agc acg aat cct aaa cct caa aga aag acc aaa cgt aac acc aac 18039  
 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn  
 1775 1780 1785

cg<sup>g</sup> cg<sup>g</sup> cc<sup>g</sup> cag gag gtc aag tt<sup>c</sup> cg<sup>g</sup> ggt ggc ggt cag atc gtt ggt 18087  
 Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gln Ile Val Gly  
 1790 1795 1800

gga gtt tac ttg ttg cc<sup>g</sup> cg<sup>c</sup> agg ggc cct aga ttg ggt gtg cg<sup>g</sup> gcg 18135  
 Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala  
 1805 1810 1815

acg aga aag act tcc gag cg<sup>g</sup> tc<sup>g</sup> c<sup>a</sup> c<sup>c</sup> t<sup>c</sup> c<sup>g</sup> a<sup>t</sup> g<sup>g</sup> t<sup>t</sup> g<sup>g</sup> t<sup>t</sup> g<sup>g</sup> 18183  
 Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro  
 1820 1825 1830 1835

atc ccc aag gct cgt cg<sup>g</sup> cc<sup>g</sup> gag ggc agg acc tgg gct cag ccc ggg 18231  
 Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly  
 1840 1845 1850

tac cct tgg ccc ctc tat ggc aat gag ggc tgc ggg tgg gcg gga tgg 18279  
 Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp  
 1855 1860 1865

ctc ctg tct ccc cgt ggc tct cg<sup>g</sup> cct agc tgg ggc ccc aca gac ccc 18327  
 Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro

1870

1875

1880

cgg cgt agg tcg cgc aat ttg ggt aag gtc atc gat acc ctt acg tgc 18375  
 Arg Arg Arg Ser Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys  
 1885

1890

1895

ggc ttc gcc gac ctc atg ggg tac ata ccg ctc gtc ggc gcc cct ctt 18423  
 Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu  
 1900 1905 1910 1915

gga ggc gct gcc agg gcc ctg gcg cat ggc gtc cgg gtt ctg gaa gac 18471  
 Gly Gly Ala Ala Arg Ala Leu Ala His Gly Val Arg Val Leu Glu Asp  
 1920 1925 1930

ggc gtg aac tat gca aca ggg aac ctt cct ggt tgc tct taatagtcga 18520  
 Gly Val Asn Tyr Ala Thr Gly Asn Leu Pro Gly Cys Ser  
 1935 1940

ctttgttccc actgtacttt tagctcgtaaaaatacatatactttca tttctccgta 18580

aacaacatgt tttccatgt aatatccttt tctattttc gttccgttac caactttaca 18640

catactttat atagctattc acttctatac actaaaaaac taagacaatt ttaattttgc 18700

tgcctgccat atttcaattt gttataaatt cctataattt atcctatttag tagctaaaaa 18760

aagatgaatg tgaatcgaat cctaagagaa ttggatctga tccacaggac ggggtggc 18820

gccatgatcg cgtatcgat agtggctcca agtagcgaag cgagcaggac tggcgccgg 18880

ccaaagcggt cggacagtgc tccgagaacg ggtgcgcata gaaattgcataa 18940

agcgctagca gcacgccata gtgactggcg atgctgtcgg aatggacgat atcccgcaag 19000

aggccccgca gtaccggcat aaccaagcct atgcctacag catccagggat gacggtgccg 19060

aggatgacga tgagcgcatt gttagatttc atacacggtg cctgactgcg ttagcaattt 19120

aactgtgata aactaccgca ttaaagcttt ttctttccaa tttttttttt ttcgtcatta 19180

taaaaatcat tacgaccgag attcccggtt aataactgat ataattaaat tgaagctcta 19240

atttgtgagt ttagtataca tgcattttact tataatacag ttttttagtt ttgctggccg 19300

catcttctca aatatgcttc ccagcctgct tttctgttaac gttcacccctc taccttagca 19360

tcccttccct ttgcaaatacg tccttttcca acaataataa tgtcagatcc tgtagagacc 19420

acatcatcca cgggtctata ctgttgaccc aatgcgtctc cttgtcatc taaaccacaca 19480

ccgggtgtca taatcaacca atcgtaacct tcatactcttc cacccatgtc tctttgagca 19540

ataaaagccga taacaaaatc ttgtcgctc ttgcataatgt caacagtacc ctttagtat 19600

tctccagtag atagggagcc cttgcataatgc aattctgtca acatcaaaag gcctcttaggt 19660

tcctttgtta cttcttctgc cgcctgcttc aaaccgctaa caatacctgg gcccaccacaca 19720

ccgtgtgcat tcgtaatgtc tgcccattct gctattctgt atacacccgc agagtactgc 19780  
aatttgaactg tattaccaat gtcagcaa at tttctgtctt cgaagagata aaaaattgtac 19840  
ttggcgata atgccttag cggttaact gtgcctcca tggaaaaatc agtcaagata 19900  
tccacatgtg ttttagtaa acaaatttg ggaccta atg cttcaactaa ctccagtaat 19960  
tccttggtgg tacgaacatc caatgaagca cacaagttt tttgctttc gtgcata 20020  
ttaaatagct tggcagcaac aggacttagga tgagtagcag cacgttcctt atatgtac 20080  
ttcgacatga ttatctcg ttcctgcag gttttgttc tgtgcagg tg 20140  
actgggcaat ttcatgttcc ttcaacacta catatgcgtatataaccatct 20200  
tgctccttcc ttcatgttcc ttcatgttcc gagattaccg aatcaaaaaa atttcaagga 20260  
aaccgaaatc aaaaaaaaaa ataaaaaaaaa aatgatgaat tgaaaagctt atcgat 20316

<210> 15  
<211> 1944  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
pd.delta.NS3NS5.pj.core173

<400> 15  
Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro  
1 5 10 15  
Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His  
20 25 30  
Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly  
35 40 45  
Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly  
50 55 60  
Cys Ser Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser  
65 70 75 80  
Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Ala  
85 90 95  
Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala Thr Ala Thr Pro Pro  
100 105 110  
Gly Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser  
115 120 125  
Thr Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val  
130 135 140

Ile Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Lys Lys Cys  
 145 150 155 160

Asp Glu Leu Ala Ala Lys Leu Val Ala Leu Gly Ile Asn Ala Val Ala  
 165 170 175

Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val  
 180 185 190

Val Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe  
 195 200 205

Asp Ser Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe  
 210 215 220

Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr Ile Thr Leu Pro Gln Asp  
 225 230 235 240

Ala Val Ser Arg Thr Gln Arg Arg Gly Arg Thr Gly Arg Gly Lys Pro  
 245 250 255

Gly Ile Tyr Arg Phe Val Ala Pro Gly Glu Arg Pro Ser Gly Met Phe  
 260 265 270

Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp Ala Gly Cys Ala Trp Tyr  
 275 280 285

Glu Leu Thr Pro Ala Glu Thr Thr Val Arg Leu Arg Ala Tyr Met Asn  
 290 295 300

Thr Pro Gly Leu Pro Val Cys Gln Asp His Leu Glu Phe Trp Glu Gly  
 305 310 315 320

Val Phe Thr Gly Leu Thr His Ile Asp Ala His Phe Leu Ser Gln Thr  
 325 330 335

Lys Gln Ser Gly Glu Asn Leu Pro Tyr Leu Val Ala Tyr Gln Ala Thr  
 340 345 350

Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser Trp Asp Gln Met Trp  
 355 360 365

Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His Gly Pro Thr Pro Leu  
 370 375 380

Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Ile Thr Leu Thr His Pro  
 385 390 395 400

Val Thr Lys Tyr Ile Met Thr Cys Met Ser Ala Asp Leu Glu Val Val  
 405 410 415

Thr Ser Thr Trp Val Leu Val Gly Gly Val Leu Ala Ala Leu Ala Ala  
 420 425 430

Tyr Cys Leu Ser Thr Gly Cys Val Val Ile Val Gly Arg Val Val Leu  
 435 440 445

Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu Val Leu Tyr Arg Glu  
 450 455 460

Phe Asp Glu Met Glu Glu Cys Ser Gln His Leu Pro Tyr Ile Glu Gln  
 465 470 475 480

Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys Ala Leu Gly Leu Leu  
 485 490 495

Gln Thr Ala Ser Arg Gln Ala Glu Val Ile Ala Pro Ala Val Gln Thr  
 500 505 510

Asn Trp Gln Lys Leu Glu Thr Phe Trp Ala Lys His Met Trp Asn Phe  
 515 520 525

Ile Ser Gly Ile Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn  
 530 535 540

Pro Ala Ile Ala Ser Leu Met Ala Phe Thr Ala Ala Val Thr Ser Pro  
 545 550 555 560

Leu Thr Thr Ser Gln Thr Leu Leu Phe Asn Ile Leu Gly Gly Trp Val  
 565 570 575

Ala Ala Gln Leu Ala Ala Pro Gly Ala Ala Thr Ala Phe Val Gly Ala  
 580 585 590

Gly Leu Ala Gly Ala Ala Ile Gly Ser Val Gly Leu Gly Lys Val Leu  
 595 600 605

Ile Asp Ile Leu Ala Gly Tyr Gly Ala Gly Val Ala Gly Ala Leu Val  
 610 615 620

Ala Phe Lys Ile Met Ser Gly Glu Val Pro Ser Thr Glu Asp Leu Val  
 625 630 635 640

Asn Leu Leu Pro Ala Ile Leu Ser Pro Gly Ala Leu Val Val Gly Val  
 645 650 655

Val Cys Ala Ala Ile Leu Arg Arg His Val Gly Pro Gly Glu Gly Ala  
 660 665 670

Val Gln Trp Met Asn Arg Leu Ile Ala Phe Ala Ser Arg Gly Asn His  
 675 680 685

Val Ser Pro Thr His Tyr Val Pro Glu Ser Asp Ala Ala Ala Arg Val  
 690 695 700

Thr Ala Ile Leu Ser Ser Leu Thr Val Thr Gln Leu Leu Arg Arg Leu  
 705 710 715 720

His Gln Trp Ile Ser Ser Glu Cys Thr Thr Pro Cys Ser Gly Ser Trp  
 725 730 735

Leu Arg Asp Ile Trp Asp Trp Ile Cys Glu Val Leu Ser Asp Phe Lys  
 740 745 750

Thr Trp Leu Lys Ala Lys Leu Met Pro Gln Leu Pro Gly Ile Pro Phe  
 755 760 765

Val Ser Cys Gln Arg Gly Tyr Lys Gly Val Trp Arg Gly Asp Gly Ile  
 770 775 780

Met His Thr Arg Cys His Cys Gly Ala Glu Ile Thr Gly His Val Lys  
 785 790 795 800

Asn Gly Thr Met Arg Ile Val Gly Pro Arg Thr Cys Arg Asn Met Trp  
 805 810 815

Ser Gly Thr Phe Pro Ile Asn Ala Tyr Thr Thr Gly Pro Cys Thr Pro  
 820 825 830

Leu Pro Ala Pro Asn Tyr Thr Phe Ala Leu Trp Arg Val Ser Ala Glu  
 835 840 845

Glu Tyr Val Glu Ile Arg Gln Val Gly Asp Phe His Tyr Val Thr Gly  
 850 855 860

Met Thr Thr Asp Asn Leu Lys Cys Pro Cys Gln Val Pro Ser Pro Glu  
 865 870 875 880

Phe Phe Thr Glu Leu Asp Gly Val Arg Leu His Arg Phe Ala Pro Pro  
 885 890 895

Cys Lys Pro Leu Leu Arg Glu Glu Val Ser Phe Arg Val Gly Leu His  
 900 905 910

Glu Tyr Pro Val Gly Ser Gln Leu Pro Cys Glu Pro Glu Pro Asp Val  
 915 920 925

Ala Val Leu Thr Ser Met Leu Thr Asp Pro Ser His Ile Thr Ala Glu  
 930 935 940

Ala Ala Gly Arg Arg Leu Ala Arg Gly Ser Pro Pro Ser Val Ala Ser  
 945 950 955 960

Ser Ser Ala Ser Gln Leu Ser Ala Pro Ser Leu Lys Ala Thr Cys Thr  
 965 970 975

Ala Asn His Asp Ser Pro Asp Ala Glu Leu Ile Glu Ala Asn Leu Leu  
 980 985 990

Trp Arg Gln Glu Met Gly Gly Asn Ile Thr Arg Val Glu Ser Glu Asn  
 995 1000 1005

Lys Val Val Ile Leu Asp Ser Phe Asp Pro Leu Val Ala Glu Glu Asp  
 1010 1015 1020

Glu Arg Glu Ile Ser Val Pro Ala Glu Ile Leu Arg Lys Ser Arg Arg  
 1025 1030 1035 1040

Phe Ala Gln Ala Leu Pro Val Trp Ala Arg Pro Asp Tyr Asn Pro Pro  
 1045 1050 1055

Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr Glu Pro Pro Val Val His  
1060 1065 1070

Gly Cys Pro Leu Pro Pro Pro Lys Ser Pro Pro Val Pro Pro Pro Arg  
1075 1080 1085

Lys Lys Arg Thr Val Val Leu Thr Glu Ser Thr Leu Ser Thr Ala Leu  
1090 1095 1100

Ala Glu Leu Ala Thr Arg Ser Phe Gly Ser Ser Ser Thr Ser Gly Ile  
105 1110 1115 1120

Thr Gly Asp Asn Thr Thr Ser Ser Glu Pro Ala Pro Ser Gly Cys  
1125 1130 1135

Pro Pro Asp Ser Asp Ala Glu Ser Tyr Ser Ser Met Pro Pro Leu Glu  
1140 1145 1150

Gly Glu Pro Gly Asp Pro Asp Leu Ser Asp Gly Ser Trp Ser Thr Val  
1155 1160 1165

Ser Ser Glu Ala Asn Ala Glu Asp Val Val Cys Cys Ser Met Ser Tyr  
1170 1175 1180

Ser Trp Thr Gly Ala Leu Val Thr Pro Cys Ala Ala Glu Glu Gln Lys  
185 1190 1195 1200

Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu Arg His His Asn Leu  
1205 1210 1215

Val Tyr Ser Thr Ser Arg Ser Ala Cys Gln Arg Gln Lys Lys Val  
1220 1225 1230

Thr Phe Asp Arg Leu Gln Val Leu Asp Ser His Tyr Gln Asp Val Leu  
1235 1240 1245

Lys Glu Val Lys Ala Ala Ser Lys Val Lys Ala Asn Leu Leu Ser  
1250 1255 1260

Val Glu Glu Ala Cys Ser Leu Thr Pro Pro His Ser Ala Lys Ser Lys  
265 1270 1275 1280

Phe Gly Tyr Gly Ala Lys Asp Val Arg Cys His Ala Arg Lys Ala Val  
1285 1290 1295

Thr His Ile Asn Ser Val Trp Lys Asp Leu Leu Glu Asp Asn Val Thr  
1300 1305 1310

Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu Val Phe Cys Val Gln  
1315 1320 1325

Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu Ile Val Phe Pro Asp  
1330 1335 1340

Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu Tyr Asp Val Val Thr  
345 1350 1355 1360

Lys Leu Pro Leu Ala Val Met Gly Ser Ser Tyr Gly Phe Gln Tyr Ser  
1365 1370 1375

Pro Gly Gln Arg Val Glu Phe Leu Val Gln Ala Trp Lys Ser Lys Lys  
1380 1385 1390

Thr Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe Asp Ser Thr Val  
1395 1400 1405

Thr Glu Ser Asp Ile Arg Thr Glu Glu Ala Ile Tyr Gln Cys Cys Asp  
1410 1415 1420

Leu Asp Pro Gln Ala Arg Val Ala Ile Lys Ser Leu Thr Glu Arg Leu  
425 1430 1435 1440

Tyr Val Gly Gly Pro Leu Thr Asn Ser Arg Gly Glu Asn Cys Gly Tyr  
1445 1450 1455

Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr Ser Cys Gly Asn Thr  
1460 1465 1470

Leu Thr Cys Tyr Ile Lys Ala Arg Ala Ala Cys Arg Ala Ala Gly Leu  
1475 1480 1485

Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp Leu Val Val Ile Cys  
1490 1495 1500

Glu Ser Ala Gly Val Gln Glu Asp Ala Ala Ser Leu Arg Ala Phe Thr  
505 1510 1515 1520

Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp Pro Pro Gln Pro  
1525 1530 1535

Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser Asn Val Ser Val  
1540 1545 1550

Ala His Asp Gly Ala Gly Lys Arg Val Tyr Tyr Leu Thr Arg Asp Pro  
1555 1560 1565

Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr Ala Arg His Thr Pro  
1570 1575 1580

Val Asn Ser Trp Leu Gly Asn Ile Ile Met Phe Ala Pro Thr Leu Trp  
585 1590 1595 1600

Ala Arg Met Ile Leu Met Thr His Phe Phe Ser Val Leu Ile Ala Arg  
1605 1610 1615

Asp Gln Leu Glu Gln Ala Leu Asp Cys Glu Ile Tyr Gly Ala Cys Tyr  
1620 1625 1630

Ser Ile Glu Pro Leu Asp Leu Pro Pro Ile Ile Gln Arg Leu His Gly  
1635 1640 1645

Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro Gly Glu Ile Asn Arg  
1650 1655 1660

Val Ala Ala Cys Leu Arg Lys Leu Gly Val Pro Pro Leu Arg Ala Trp  
 665 1670 1675 1680  
 Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu Leu Ala Arg Gly Gly  
 1685 1690 1695  
 Arg Ala Ala Ile Cys Gly Lys Tyr Leu Phe Asn Trp Ala Val Arg Thr  
 1700 1705 1710  
 Lys Leu Lys Leu Thr Pro Ile Ala Ala Gly Gln Leu Asp Leu Ser  
 1715 1720 1725  
 Gly Trp Phe Thr Ala Gly Tyr Ser Gly Gly Asp Ile Tyr His Ser Val  
 1730 1735 1740  
 Ser His Ala Arg Pro Arg Trp Ile Trp Phe Cys Leu Leu Leu Ala  
 745 1750 1755 1760  
 Ala Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg Met Ser Thr Asn Pro  
 1765 1770 1775  
 Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Gln Asp  
 1780 1785 1790  
 Val Lys Phe Pro Gly Gly Gln Ile Val Gly Gly Val Tyr Leu Leu  
 1795 1800 1805  
 Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala Thr Arg Lys Thr Ser  
 1810 1815 1820  
 Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro Ile Pro Lys Ala Arg  
 825 1830 1835 1840  
 Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly Tyr Pro Trp Pro Leu  
 1845 1850 1855  
 Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp Leu Leu Ser Pro Arg  
 1860 1865 1870  
 Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro Arg Arg Ser Arg  
 1875 1880 1885  
 Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu  
 1890 1895 1900  
 Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg  
 905 1910 1915 1920  
 Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala  
 1925 1930 1935  
 Thr Gly Asn Leu Pro Gly Cys Ser  
 1940

<210> 16  
 <211> 20217

<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
pd.delta.NS3NS5.pj.core140

<220>  
<221> CDS  
<222> (12679) . . (18411)

ggtttgcattt gattggaaat ggtatgctct cttcacaaga agcaagtcag gctgccatag 1440  
atttgatgct acagaataac aagctgttag acaatagaaaa gcaactatac aaatctattg 1500  
ctataataat aggaagattt cccgagaaaag acaagaagag agctaccgaa atgctcatga 1560  
gaaaaatgga ttgtacacag ttattagtcc caccagctcc aacggaagaa gatgttatga 1620  
agctcgtaag cgtcgttacc caattgctta cttagttcc accagatcgtaaagctgctt 1680  
taataggta tttattcattc ccggaatctc taaaggatatttcaatgaac 1740  
tggcggcaga gaatcgttt cagcaaaaaa agagttagt ggaaggaagg actgaagtga 1800  
accatgctaa tacaaatgaa gaagttccct ccaggcgaac aagaagtata gacacaaatg 1860  
caagaggagc atataaatta caaaacacca tcactgaggg ccctaaagcg gttcccacga 1920  
aaaaaaggag agtagcaacg aggtaaggg gcagaaaaatc acgttaatact tctagggtat 1980  
gatccaatat caaaggaaat gatagcattt aaggatgaga ctaatocaat tgaggagtgg 2040  
cagcatatac aacagctaaa ggtagtgctt gaaggaagca tacgataccc cgcatggaat 2100  
gggataatat cacaggaggt actagactac cttagtccat acataaatac acgcatataa 2160  
gtacgcattt aagcataaac acgcactatg ccgttcttctt catgtatata tatatacagg 2220  
caacacgcag atatagggtc gacgtgaaca gtgagctgtt tgcgcgcagc tcgcgttgca 2280  
tttcggaaag cgctcgaaaaa cggaaacgct ttgaagttcc tattccgaag ttcctattct 2340  
ctagaaagta taggaacttc agagcgcttt tgaaaaccaa aagcgctctg aagacgcact 2400  
ttcaaaaaac caaaaacgca ccggactgtt acgagctactaaaatattgc gaataccgct 2460  
tccacaaaca ttgctaaaaa gtatctttt gctatatac tctgtgttat atccctatata 2520  
aacctaccca tccacccccc gtccttgcattt cttgcataaactcgacct ctacatcaac 2580  
aggcttccaa tgcttttcaaa attttactgtt caagtagacc catacggtt taatatgtt 2640  
ctcttcataa tgtaagctta tctttatcgaa atcgtgtgaa aaactactac cgcatgaaac 2700  
ctttacgggtt ccctgagattt gaatttagttt cttagtata tgatacaaga cactttgaa 2760  
ctttgtacga cgaattttga gtttcggccat cctctggctt tttccaaatata tcctgtcgcc 2820  
tattatctcc gcctcagttt gatctccgc ttccagactgc cattttcac ataatgaatc 2880  
tatttcaccc cacaatcctt catccgcctt cgcattttgtt tccgttaaac tattgacttc 2940  
atgttgtaca ttgttttagttt cacgagaagg gtcctttca ggcggtagctt cctgatctcc 3000  
tatatgaccc ttatcctgtt ctcttccac aaacttagaa atgtattcat gaattatgga 3060  
gcacctaata acattttca aggccggagaa gtttggccaa gatgcccaat atgcttgaca 3120

DRAFT FEB 2000

tgaaaacgtg agaatgaatt tagtattatt gtgatattct gaggcaattt tattataatc 3180  
tcgaagataa gagaagaatg cagtgacctt tgtattgaca aatggagatt ccatgtatct 3240  
aaaaaaatacg cctttaggcc ttctgatacc ctccccctg cggttagcg tgccctttac 3300  
attaatatct aaaccctctc cgatggtggc cttaactga ctaataaaatg caaccgatat 3360  
aaactgtat aattctgggt gatTTATGAT tcgatcgaca attgtattgt acactagtgc 3420  
aggatcagggc caatccagtt ctTTTcaat taccgggtgt tcgtctgtat tcagtgatg 3480  
tccaacaaat gcaaATGCTA acgtttgtt tttcttataa ttgtcaggaa ctggaaaagt 3540  
cccccttgcgtc gtctcgatta cacacctact ttcatcgta accatagggtt ggaagtgctg 3600  
cataatacat tgcttaatac aagcaagcag tctctcgcca ttcatatttc agttattttc 3660  
cattacagct gatgtcattt tatATCAGCG ctgtaaaaat ctatctgtta cagaaggTTT 3720  
tcgcggTTT tataaacaaa actttcgTTA cgaaatcgag caatcacccc agctgcgtat 3780  
ttggaaattc gggaaaaagt agagcaacgc gagttgcatt ttTACACCCA taatgcata 3840  
ttaacttcga gaagggatta aggctaattt cactagtatg ttcaaaaac ctcaatctgt 3900  
ccattgaatg ccttataaaa cagctataga ttgcataaaaa gagttagcta ctcaatgctt 3960  
tttgtcaaaag cttaactgtatg atgatgtgtc tactttcagg cgggtctgtt gtaaggagaa 4020  
tgacattata aagctggcac tttagaattcc acggactata gactatacta gtatactccg 4080  
tctactgtac gatacacttc cgctcaggTC cttgtcttt aacgaggcct taccacttt 4140  
ttgttactct attgatccag ctcagcaag gcaggtgtat ctaagattct atttcgcga 4200  
tgttagaaaa ctagctagac cgagaaagag actagaaatg caaaaggcac ttctacaatg 4260  
gctgccatca ttattatccg atgtgacgct gatTTTTTT tttttttttt tttttttttt 4320  
ttttttttt tttttttttt tttttggta caaatatcat aaaaaaaagag aatctttta 4380  
agcaaggatt ttcttaactt ctccggcgc agcatcacccg acttcgggtgg tactgttgaa 4440  
accacccaaa tcaccagttc tgatacctgc atccaaaacc ttttaactg catcttcaat 4500  
ggcttaccc tcttcaggca agttcaatgca caatttcaac atcattgcag cagacaagat 4560  
agtggcgata gggttgaccc tatttttgg caaatctgga gcggaaaccat ggcattggc 4620  
gtacaaacca aatgcgggtgt tctgtctgg caaagaggcc aaggacgcag atggcaacaa 4680  
acccaaggag cctggataa cggaggcttc atcggagatg atatcaccaaa acatgttgct 4740  
ggtgattata ataccattta gttgggttgg gttcttaact aggatcatgg cggcagaatc 4800  
aatcaattga tttttttttt tcaatgttagg gaattcgttc ttgtatggTTT cctccacagt 4860

tttctccat aatcttgaag aggccaaaac attagctta tccaaggacc aaataggcaa 4920  
tggggctca tggttaggg ccatgaaagc ggccattctt gtgattctt gcacttctgg 4980  
aacggtgtat tgttcaatat cccaagcgac accatcacca tcgtcttct ttctcttacc 5040  
aaagtaaata cctcccacta attctctaac aacaacgaag tcagtacctt tagcaaattg 5100  
tggcttgatt ggagataagt ctAAAAGAGA gtcggatgca aagttacatg gtTTAAAGTT 5160  
ggcgtacaat tgaagttctt tacggatttt tagtaaacct tggcggatgc taacactacc 5220  
ggtaccccat ttaggaccac ccacagcacc taacaaaacg gcatcagcct tcttggaggc 5280  
ttccagcgcc tcatactggaa gtggAACACC tggatcgatcg atagcagcAC caccaattaa 5340  
atgatTTCG aaatcgaact tgacattgga acgaacatca gaaatAGCTT taagaacctt 5400  
aatggcttcg gctgtgattt cttgaccaac gtggcacct ggcaaaacgA cgatcttctt 5460  
aggggcagac attacaatgg tatATCCTTG aaatataat aaaaaaaaaaaaaaaa 5520  
aaaaaaaaaa atgcagcttc tcaatgatAT tcgaatacgc tttgaggaga tacagcctaa 5580  
tatccgacaa actgtttac agattacga tcgtacttgt taCCCATCAT tgaatTTGA 5640  
acatccgAAC ctgggagTTT tccCTGAAAC agatAGTATA tttGAACCTG tataataata 5700  
tatagtctag cgcttacgg aagacaatgt atgtatTCG tttcctggag aaactattgc 5760  
atctattgca taggtaatct tgcacgtcgc atccccggTT cattttctgc gtttccatct 5820  
tgcacttcaa tagcatatct ttgttaacga agcatctgtg cttcattttg tagaacaAAA 5880  
atgcaacgcg agagcgctaa ttttcaaac aaagaatGT agctgcATT ttacagaaca 5940  
gaaatgcaac gcgaaAGCGC tattttacca acgaAGAATC tggatTTCA ttttGAAAAA 6000  
caaaaatgca acgcgagAGC gctaattttt caaaACAAGA atctgagctg catttttaca 6060  
gaacagaaat gcaacgcgag agcgctattt tACCAACAAA gaatctatac ttcttttttG 6120  
ttctacaaaaa atgcatcccc agagcgctat ttttctaaca aagcatcttA gattactttt 6180  
tttctccTTT gtgcgctcta taatgcagtc tttgataac ttttgcact gtaggtccgt 6240  
taaggTTAGA agaaggctac tttggTGTCT attttctttt ccataaaaaaa agcctgactc 6300  
cacttcccgc gtttactgat tactagcAA gctgcgggtg catttttca agataaaggc 6360  
atccccgatt atattctata ccgatgtgga ttgcgcatac tttgtgaaca gaaagtgata 6420  
gcgttgatga ttcttcattt gtcagAAAAT tatgaacggt ttcttctatt ttgtcttat 6480  
atactacgta taggaaatgt ttacatttc gtattgtttt cgattcactc tatgaatagt 6540  
tcttactaca attttttGt ctaaAGAGTA atactagaga taaacataaa aaatgttagag 6600

gtcgagttt aatgcaggaa aaggagcga aaggatggatg ggttaggtat atagggat 6660  
agcacagaga tatatacgaa agagatactt ttgagcaatg tttgtgaaag cggtattcgc 6720  
aatatTTAG tagctcgta cagtccggtg cgTTTTGGT ttttgaaag tcgtcttca 6780  
gagcgcttt ggTTTCAAA agcgctctga agttcctata ctttcttagag aataggaact 6840  
tcggaatagg aacttcaaag cgTTTCCGAA aacgagcgct tccgaaaatg caacgcgagc 6900  
tgcgcacata cagctcactg ttcacgtcgc acctatatct gcgtgttgc tgtatata 6960  
tatacatgag aagaacggca tagtgcgtgt ttatgcttaa atgcgtactt atatgcgtct 7020  
atttatgttag gatgaaaggt agtctagtagc ctccgtgat attatccat tccatgcggg 7080  
gtatcgatg cttccTTcag cactaccctt tagctgttct atatgcgtcc actcctcaat 7140  
tggatttagtc tcattcTTca atgctatcat ttccTTgat attggatcat atgcata 7200  
ccgagaaact agtgcgaagt agtgcgttctt atctgatgag tatacgTTgt 7260  
cctggccacg gcagaagcac gcttatcgct ccaatttccc acaacattag tcaactccgt 7320  
taggccCTTC attgaaagaa atgaggtcat caaatgtctt ccaatgtgag attttggcc 7380  
atTTTTATA gcaaagattg aataaggcgc atTTTCTTC aaagctttat tgtacgtct 7440  
gactaagtta tctttataa attggatttc ctgttattg cttgaagaat tgccggcct 7500  
atTTTACTCGT tttaggactg gttcagaatt cctcaaaaat tcattcaaat atacaagtgg 7560  
atcgatgata agctgtcaaa catgagaatt cttgaagacg aaaggccctc gtgatacgcc 7620  
tatTTTTATA ggttaatgtc atgataataa tggTTTCTTA gacgtcaggt ggcactttc 7680  
ggggaaatgt gcgcggAACc CCTATTGTT TATTTCTA aatacattca aatatgtatc 7740  
cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg aagagtatga 7800  
gtattcaaca ttccgtgtc gccCTTATTc CCTTTTGC ggcattttgc CTTCTGTtT 7860  
ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag 7920  
tggTTTACAT cgaactggat ctcaacagcg gtaagatcct tgagagttt cgccccgaag 7980  
aacgtttcc aatgatgagc actTTTAAAG ttctgctatg tggcgccgta ttatcccgt 8040  
ttgacGCCGG gcaagagcaa ctcggTcGCC gcatacacta ttctcagaat gacttgggt 8100  
agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca 8160  
gtgctGCCat aaccatgagt gataacactg cggccaactt acttctgaca acgatcgag 8220  
gaccgaagga gctaaccgct ttttgcaca acatggggga tcatgtaact cgccttgatc 8280  
gttgggaacc ggagctgaat gaagccatac caaAcgcga gCgtgacacc acgatgcctg 8340

cagcaatggc aacaacgttg cgcaaactat taactggcg aactacttact ctagttccc 8400  
ggcaacaatt aatagactgg atggaggcg aataaagtgc aggaccactt ctgcgctcg 8460  
cccttcggc tggctggttt attgctgata aatctgggc cggtgagcgt gggtcgcg 8520  
gtatcattgc agcaactgggg ccagatggta agccctcccg tatacgtagtt atcacacga 8580  
cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata ggcgcctcac 8640  
tgattaagca ttggtaactg tcagaccaag tttactcata tatactttag attgatttaa 8700  
aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttataat ctcatgacca 8760  
aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag 8820  
gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaaca aaaaaaccac 8880  
cgctaccagc ggtggttgt ttgccggatc aagagctacc aactctttt ccgaaggtaa 8940  
ctggcttcag cagagcgcag ataccaaata ctgtccttct agtgtagccg tagttaggcc 9000  
accacttcaa gaactctgta gcaccgccta catacctcgc tctgtaatc ctgttaccag 9060  
tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac 9120  
cggataaggc gcagcggcgcg ggctgaacgg ggggttcgtg cacacagccc agcttggagc 9180  
gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc 9240  
ccgaaggggag aaaggcggac aggtatccgg taagcggcag ggtcggaaaca ggagagcgca 9300  
cgagggagct tccaggggaa aacgcctggt atctttatag tccctgtcggg tttcgccacc 9360  
tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg 9420  
ccagcaacgc ggcctttta cggttcctgg cctttgctg gcctttgct cacatgttct 9480  
ttcctgcgtt atccccgtat tctgtggata accgttattac cgccttgag tgagctgata 9540  
ccgctcgccg cagccgaacg accgagcgcg gcgagtcagt gagcgaggaa gcggaaagagc 9600  
gcctgatgcg gtattttctc cttacgcattc tgtcggtat ttcacaccgc atatggcga 9660  
ctctcagtac aatctgctct gatgcccatc agttaagcca gtatacactc cgctatcgct 9720  
acgtgactgg gtcatggctg cggccgaca cccgccaaca cccgctgacg cgccctgacg 9780  
ggcttgcgtc cttccggcat cggcttadag acaagctgtc accgtctccg ggagctgcatt 9840  
gtgtcagagg ttttacccgt catcacccaa acgcgcgagg cagctgcgtt aaagctcatc 9900  
agcgtggcgcg tgaagcgatt cacagatgtc tgcctgttca tccgcgttcca gctcggtt 9960  
tttctccaga agcgttaatg tctggcttct gataaagcgg gccatgtttaa gggcggtttt 10020  
ttcctgtttt gtcactgatc cttccgttta agggggattt ctgttcatgg gggtaatgt 10080

accgatgaaa cgagagagga tgctcacat acgggtact gatgtgaac atgcccggtt 10140  
actggAACGT tggaggta aacaactggc ggtatggatg cggcgggacc agagaaaaat 10200  
caactcagggt caatgccagc gcttcgttaa tacagatgtt ggtgttccac aggttagcca 10260  
gcagcatcct gcgtgcaga tccggAACAT aatggtgca ggcgctgact tccgcgtttc 10320  
cagactttac gaaacacgga aaccgaagac cattcatgtt gttgctcagg tcgcagacgt 10380  
tttgcagcag cagtgcgttc acgttcgctc gcgtatcggt gattcattct gctaaccagt 10440  
aaggcaaccc cgccagccta gccgggtcct caacgacagg agcacgatca tgccacccg 10500  
tggccaggac ccaacgctgc ccgagatgca cgcgtgcgg ctgctggaga tggcggacgc 10560  
gatggatatg ttctgccaag gttgggttg cgcattcaca gttctccgca agaattgatt 10620  
ggctccaatt cttggagtgg tgaatccgtt agcgaggtgc cgccgggttc cattcaggtc 10680  
gaggtggccc ggctccatgc accgcgacgc aacgcgggaa ggcagacaag gtatagggcg 10740  
gcgcctacaa tccatgcaa cccgttccat gtgctgcgg aggoggcata aatgcgcgtg 10800  
acgatcagcg gtccaatgat cgaagttagg ctggtaagag cgcgcgacgc tccttgaagc 10860  
tgtccctgat ggtcgcatc tacctgcctg gacagcatgg cctgcaacgc gggcatcccg 10920  
atgcccgg aagcgagaag aatcataatg gggaaaggcca tccagcctcg cgtcgcgaac 10980  
gccagcaaga cgtagcccg cgcgtcgccc gccatgcgg cgataatggc ctgcttctcg 11040  
ccgaaacgtt tggggcggg accagtgcacg aaggcttgg cgagggcgtg caagattccg 11100  
aataccgcaa ggcacaggcc gatcatcgac ggcgcgtc gaaagcggtc ctcgcggaaa 11160  
atgacccaga ggcgtgcgg cacctgtcct acgagttgca tgataaagaa gacagtcata 11220  
agtgcggcga cgatagtcat gccccgcggc cacggaaagg agctgactgg gttgaaggct 11280  
ctcaaggcga tcggtcgagg atccttcaat atgcgcacat acgctgttat gttcaaggtc 11340  
ccttcgttta agaacgaaag cggcttcct tttgaggat gttcaagtt gttcaatct 11400  
atcaaatttg caaatccccca gtctgtatct agagcggtga atcggtgatg cgatttggta 11460  
attaaattga tgggtgcacc attaccaggc ctagatatac caatggcaaa ctgagcacaa 11520  
caataccagt ccggatcaac tggcaccatc tctccgtag tctcatctaa ttttcttcc 11580  
ggatgagggtt ccagatatac cgcaacaccc ttattatggt ttccctgagg gaataataga 11640  
atgtcccatt cgaaatcacc aattccaaac ctggcgaat tgtatccgg gtttggtaac 11700  
tcgttccagt caggaatgtt ccacgtgaag ctatcttcca gcaaagtctc cacttcttca 11760  
tcaaattgtg gagaatactc ccaatgctct tatctatggg acttccggaa aacacagtac 11820

cgataacttcc caattcgctc tcagagctca ttgtttgttt gaagagacta atcaaagaat 11880  
 cgaaaaattaa tatcttaact gatagtttga tcaaaggggc aaaacgtagg 11940  
 ggcaaaacaaa cggaaaaatc gtttctcaaa ttttctgatg ccaagaactc taaccagtct 12000  
 tatctaaaaa ttgccttatg atccgtctct ccgggtacag cctgtgtaac tgattaatcc 12060  
 tgccttcta atcaccattc taatgttta attaaggat tttgtcttca ttaacggctt 12120  
 tcgctcataa aaatgttatg acgaaaagcc cgcaggcgaa aaaccatcca cttcacgaga 12180  
 ctgatctcct ctgcccgaac accgggcattc tccaaattat aagttggaga aataagagaa 12240  
 tttcagattt agagaatgaa aaaaaaaaaac ctttagttca taggtccatt ctcttagcgc 12300  
 aactacagag aacagggca caaacaggca aaaaacggc acaacatcaa tggagtgtatg 12360  
 caacctgcct ggagtaaatg atgacacaag gcaattgacc cacgtatgtatcat 12420  
 tttcttacac cttcttattttt cttctgctct ctctgatttg gaaaaagctg aaaaaaaaaagg 12480  
 ttgaaaccag ttccctgaaa ttattccctt acttgactaa taagtatata aagacggtag 12540  
 gtattgattt taattctgtatc aatctatttc ttaaacttct taaattctac tttttagtt 12600  
 agtctttttt ttagttttaa aacaccaaga acttagtttca gaataaacac acataaaacaa 12660  
 acaagcttac aaaacaaa atg gct gca tat gca gct cag ggc tat aag gtg 12711  
                   Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val  
                   1              5                  10  
  
 cta gta ctc aac ccc tct gtt gct gca aca ctg ggc ttt ggt gct tac 12759  
 Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr  
                   15              20                  25  
  
 atg tcc aag gct cat ggg atc gat cct aac atc agg acc ggg gtg aga 12807  
 Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg  
                   30              35                  40  
  
 aca att acc act ggc agc ccc atc aac tac tcc acc tac ggc aag ttc 12855  
 Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe  
                   45              50                  55  
  
 ctt gcc gac ggc ggg tgc tcg ggg ggc gct tat gac ata ata att tgt 12903  
 Leu Ala Asp Gly Gly Cys Ser Gly Ala Tyr Asp Ile Ile Ile Cys  
                   60              65                  70                  75  
  
 gac gag tgc cac tcc acg gat gcc aca tcc atc ttg ggc att ggc act 12951  
 Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr  
                   80              85                  90  
  
 gtc ctt gac caa gca gag act gcg ggg ggc aga ctg gtt gtg ctc gcc 12999  
 Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala  
                   95              100                  105  
  
 acc gcc acc cct ccg ggc tcc gtc act gtg ccc cat ccc aac atc gag 13047

Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn Ile Glu			
110	115	120	
gag gtt gct ctg tcc acc acc gga gag atc cct ttt tac ggc aag gct		13095	
Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala			
125	130	135	
atc ccc ctc gaa gta atc aag ggg ggg aga cat ctc atc ttc tgt cat		13143	
Ile Pro Leu Glu Val Ile Lys Gly Arg His Leu Ile Phe Cys His			
140	145	150	155
tca aag aag aag tgc gac gaa ctc gcc gca aag ctg gtc gca ttg ggc		13191	
Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Ala Leu Gly			
160	165	170	
atc aat gcc gtg gcc tac tac cgc ggt ctt gac gtg tcc gtc atc ccg		13239	
Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile Pro			
175	180	185	
acc agc ggc gat gtt gtc gtc gtc gca acc gat gcc ctc atg acc ggc		13287	
Thr Ser Gly Asp Val Val Val Ala Thr Asp Ala Leu Met Thr Gly			
190	195	200	
tat acc ggc gac ttc gac tcg gtg ata gac tgc tat acg tgt gtc acc		13335	
Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys Val Thr			
205	210	215	
cag aca gtc gat ttc agc ctt gac cct acc ttc acc att gag aca atc		13383	
Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr Ile			
220	225	230	235
acg ctc ccc caa gat gct gtc tcc cgc act caa cgt cgg ggc agg act		13431	
Thr Leu Pro Gln Asp Ala Val Ser Arg Thr Gln Arg Arg Gly Arg Thr			
240	245	250	
ggc agg ggg aag cca ggc atc tac aga ttt gtg gca ccg ggg gag cgc		13479	
Gly Arg Gly Lys Pro Gly Ile Tyr Arg Phe Val Ala Pro Gly Glu Arg			
255	260	265	
ccc tcc ggc atg ttc gac tcg tcc gtc ctc tgt gag tgc tat gac gca		13527	
Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp Ala			
270	275	280	
ggc tgt gct tgg tat gag ctc acg ccc gcc gag act aca gtt agg cta		13575	
Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Thr Val Arg Leu			
285	290	295	
cga gcg tac atg aac acc ccg ggg ctt ccc gtg tgc cag gac cat ctt		13623	
Arg Ala Tyr Met Asn Thr Pro Gly Leu Pro Val Cys Gln Asp His Leu			
300	305	310	315
gaa ttt tgg gag ggc gtc ttt aca ggc ctc act cat ata gat gcc cac		13671	
Glu Phe Trp Glu Gly Val Phe Thr Gly Leu Thr His Ile Asp Ala His			
320	325	330	
ttt cta tcc cag aca aag cag agt ggg gag aac ctt cct tac ctg gta		13719	
Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Leu Pro Tyr Leu Val			

335	340	345		
gct tac caa gcc acc gtg tgc gct agg gct caa gcc cct ccc cca tcg Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser			13767	
350	355	360		
tgg gac cag atg tgg aag tgt ttg att cgc ctc aag ccc acc ctc cat Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His			13815	
365	370	375		
ggg cca aca ccc ctg cta tac aga ctg ggc gct gtt cag aat gaa atc Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Ile			13863	
380	385	390	395	
acc ctg acg cac cca gtc acc aaa tac atc atg aca tgc atg tcg gcc Thr Leu Thr His Pro Val Thr Lys Tyr Ile Met Thr Cys Met Ser Ala	400	405	410	13911
gac ctg gag gtc acg agc acc tgg gtg ctc gtt ggc ggc gtc ctg Asp Leu Glu Val Val Thr Ser Thr Trp Val Leu Val Gly Gly Val Leu	415	420	425	13959
430	435	440		
gct gct ttg gcc gcg tat tgc ctg tca aca ggc tgc gtg gtc ata gtg Ala Ala Leu Ala Ala Tyr Cys Leu Ser Thr Gly Cys Val Val Ile Val			14007	
ggc agg gtc gtc ttg tcc ggg aag ccg gca atc ata cct gac agg gaa Gly Arg Val Val Leu Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu	445	450	455	14055
460	465	470	475	
gtc ctc tac cga gag ttc gat gag atg gaa gag tgc tct cag cac tta Val Leu Tyr Arg Glu Phe Asp Glu Met Glu Cys Ser Gln His Leu				14103
ccg tac atc gag caa ggg atg atg ctc gcc gag cag ttc aag cag aag Pro Tyr Ile Glu Gln Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys	480	485	490	14151
495	500	505		
gcc ctc ggc ctc ctg cag acc gcg tcc cgt cag gca gag gtt atc gcc Ala Leu Gly Leu Leu Gln Thr Ala Ser Arg Gln Ala Glu Val Ile Ala				14199
510	515	520		
cct gct gtc cag acc aac tgg caa aaa ctc gag acc ttc tgg gcg aag Pro Ala Val Gln Thr Asn Trp Gln Lys Leu Glu Thr Phe Trp Ala Lys				14247
525	530	535		
cat atg tgg aac ttc atc agt ggg ata caa tac ttg gcg ggc ttg tca His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala Gly Leu Ser				14295
540	545	550	555	
acg ctg cct ggt aac ccc gcc att gct tca ttg atg gct ttt aca gct Thr Leu Pro Gly Asn Pro Ala Ile Ala Ser Leu Met Ala Phe Thr Ala				14343
560				
gct gtc acc agc cca cta acc act agc caa acc ctc ctc ttc aac ata Ala Val Thr Ser Pro Leu Thr Thr Ser Gln Thr Leu Leu Phe Asn Ile	565	570		14391

▲

ttg ggg ggg tgg gtg gct gcc cag ctc gcc gcc ccc ggt gcc gct act	14439
Leu Gly Gly Trp Val Ala Ala Gln Leu Ala Ala Pro Gly Ala Ala Thr	
575 580 585	
 gcc ttt gtg ggc gct ggc tta gct ggc gcc atc ggc agt gtt gga	14487
Ala Phe Val Gly Ala Gly Leu Ala Gly Ala Ala Ile Gly Ser Val Gly	
590 595 600	
 ctg ggg aag gtc ctc ata gac atc ctt gca ggg tat ggc gcg ggc gtg	14535
Leu Gly Lys Val Leu Ile Asp Ile Leu Ala Gly Tyr Gly Ala Gly Val	
605 610 615	
 gcg gga gct ctt gtg gca ttc aag atc atg agc ggt gag gtc ccc tcc	14583
Ala Gly Ala Leu Val Ala Phe Lys Ile Met Ser Gly Glu Val Pro Ser	
620 625 630 635	
 acg gag gac ctg gtc aat cta ctg ccc gcc atc ctc tcg ccc gga gcc	14631
Thr Glu Asp Leu Val Asn Leu Leu Pro Ala Ile Leu Ser Pro Gly Ala	
640 645 650	
 ctc gta gtc ggc gtg gtc tgt gca gca ata ctg cgc cgg cac gtt ggc	14679
Leu Val Val Gly Val Val Cys Ala Ala Ile Leu Arg Arg His Val Gly	
655 660 665	
 ccg ggc gag ggg gca gtg cag tgg atg aac cgg ctg ata gcc ttc gcc	14727
Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile Ala Phe Ala	
670 675 680	
 tcc cgg ggg aac cat gtt tcc ccc acg cac tac gtg ccg gag agc gat	14775
Ser Arg Gly Asn His Val Ser Pro Thr His Tyr Val Pro Glu Ser Asp	
685 690 695	
 gca gct gcc cgc gtc act gcc ata ctc agc agc ctc act gta acc cag	14823
Ala Ala Ala Arg Val Thr Ala Ile Leu Ser Ser Leu Thr Val Thr Gln	
700 705 710 715	
 ctc ctg agg cga ctg cac cag tgg ata agc tcg gag tgt acc act cca	14871
Leu Leu Arg Arg Leu His Gln Trp Ile Ser Ser Glu Cys Thr Thr Pro	
720 725 730	
 tgc tcc ggt tcc tgg cta agg gac atc tgg gac tgg ata tgc gag gtg	14919
Cys Ser Gly Ser Trp Leu Arg Asp Ile Trp Asp Trp Ile Cys Glu Val	
735 740 745	
 ttg agc gac ttt aag acc tgg cta aaa gct aag ctc atg cca cag ctg	14967
Leu Ser Asp Phe Lys Thr Trp Leu Lys Ala Lys Leu Met Pro Gln Leu	
750 755 760	
 cct ggg atc ccc ttt gtg tcc tgc cag cgc ggg tat aag ggg gtc tgg	15015
Pro Gly Ile Pro Phe Val Ser Cys Gln Arg Gly Tyr Lys Gly Val Trp	
765 770 775	
 cga ggg gac ggc atc atg cac act cgc tgc cac tgt gga gct gag atc	15063
Arg Gly Asp Gly Ile Met His Thr Arg Cys His Cys Gly Ala Glu Ile	
780 785 790 795	
 act gga cat gtc aaa aac ggg acg atg agg atc gtc ggt cct agg acc	15111

▼

卷之三

1020	1025	1030	1035	
cg <del>g</del> aag tct cg <del>g</del> aga ttc gcc cag gcc ctg ccc gtt tgg g <del>c</del> g cg <del>g</del> cc <del>g</del>				15831
Arg Lys Ser Arg Arg Phe Ala Gln Ala Leu Pro Val Trp Ala Arg Pro				
1040	1045	1050		
gac tat aac ccc cc <del>g</del> cta gtg gag ac <del>g</del> tgg aaa aag ccc gac tac gaa				15879
Asp Tyr Asn Pro Pro Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr Glu				
1055	1060	1065		
cca cct gtg gtc cat ggc tgc cc <del>g</del> ctt cca cct cca aag tcc cct cct				15927
Pro Pro Val Val His Gly Cys Pro Leu Pro Pro Pro Lys Ser Pro Pro				
1070	1075	1080		
gtg cct cc <del>g</del> cct cc <del>g</del> aag aag cg <del>g</del> ac <del>g</del> gtg gtc ctc act gaa tca acc				15975
Val Pro Pro Pro Arg Lys Lys Arg Thr Val Val Leu Thr Glu Ser Thr				
1085	1090	1095		
cta tct act gcc ttg gcc gag ctc gcc acc aga agc ttt ggc agc tcc				16023
Leu Ser Thr Ala Leu Ala Glu Leu Ala Thr Arg Ser Phe Gly Ser Ser				
1100	1105	1110	1115	
tca act tcc ggc att ac <del>g</del> ggc gac aat ac <del>g</del> aca aca tcc tct gag ccc				16071
Ser Thr Ser Gly Ile Thr Gly Asp Asn Thr Thr Ser Ser Glu Pro				
1120	1125	1130		
gcc cct tct ggc tgc ccc ccc gac tcc gac gct gag tcc tat tcc tcc				16119
Ala Pro Ser Gly Cys Pro Pro Asp Ser Asp Ala Glu Ser Tyr Ser Ser				
1135	1140	1145		
atg ccc ccc ctg gag ggg gag cct ggg gat cc <del>g</del> gat ctt agc gac ggg				16167
Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu Ser Asp Gly				
1150	1155	1160		
tca tgg tca ac <del>g</del> gtc agt agt gag gcc aac g <del>c</del> g gag gat gtc gtg tgc				16215
Ser Trp Ser Thr Val Ser Ser Glu Ala Asn Ala Glu Asp Val Val Cys				
1165	1170	1175		
tgc tca atg tct tac tct tgg aca gg <del>c</del> gca ctc gtc acc cc <del>g</del> tgc gcc				16263
Cys Ser Met Ser Tyr Ser Trp Thr Gly Ala Leu Val Thr Pro Cys Ala				
1180	1185	1190	1195	
gc <del>g</del> gaa gaa cag aaa ctg ccc atc aat gca cta agc aac tc <del>g</del> ttg cta				16311
Ala Glu Glu Gln Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu				
1200	1205	1210		
cgt cac cac aat ttg gtg tat tcc acc acc tca cgc agt gct tgc caa				16359
Arg His His Asn Leu Val Tyr Ser Thr Ser Arg Ser Ala Cys Gln				
1215	1220	1225		
agg cag aag aaa gtc aca ttt gac aga ctg caa gtt ctg gac agc cat				16407
Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Ser His				
1230	1235	1240		
tac cag gac gta ctc aag gag gtt aaa gca g <del>c</del> g g <del>c</del> tca aaa gtg aag				16455
Tyr Gln Asp Val Leu Lys Glu Val Lys Ala Ala Ser Lys Val Lys				
1245	1250	1255		

gct aac ttg cta tcc gta gag gaa gct tgc agc ctg acg ccc cca cac Ala Asn Leu Leu Ser Val Glu Glu Ala Cys Ser Leu Thr Pro Pro His 1260 1265 1270 1275	16503
tca gcc aaa tcc aag ttt ggt tat ggg gca aaa gac gtc cgt tgc cat Ser Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Cys His 1280 1285 1290	16551
gcc aga aag gcc gta acc cac atc aac tcc gtg tgg aaa gac ctt ctg Ala Arg Lys Ala Val Thr His Ile Asn Ser Val Trp Lys Asp Leu Leu 1295 1300 1305	16599
gaa gac aat gta aca cca ata gac act acc atc atg gct aag aac gag Glu Asp Asn Val Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu 1310 1315 1320	16647
gtt ttc tgc gtt cag cct gag aag ggg ggt cgt aag cca gct cgt ctc Val Phe Cys Val Gln Pro Glu Lys Gly Arg Lys Pro Ala Arg Leu 1325 1330 1335	16695
atc gtg ttc ccc gat ctg ggc gtg cgc gtg tgc gaa aag atg gct ttg Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu 1340 1345 1350 1355	16743
tac gac gtg gtt aca aag ctc ccc ttg gcc gtg atg gga agc tcc tac Tyr Asp Val Val Thr Lys Leu Pro Leu Ala Val Met Gly Ser Ser Tyr 1360 1365 1370	16791
gga ttc caa tac tca cca gga cag cgg gtt gaa ttc ctc gtg caa gcg Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Gln Ala 1375 1380 1385	16839
tgg aag tcc aag aaa acc cca atg ggg ttc tcg tat gat acc cgc tgc Trp Lys Ser Lys Thr Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys 1390 1395 1400	16887
ttt gac tcc aca gtc act gag agc gac atc cgt acg gag gag gca atc Phe Asp Ser Thr Val Thr Glu Ser Asp Ile Arg Thr Glu Glu Ala Ile 1405 1410 1415	16935
tac caa tgt tgt gac ctc gac ccc caa gcc cgc gtg gcc atc aag tcc Tyr Gln Cys Cys Asp Leu Asp Pro Gln Ala Arg Val Ala Ile Lys Ser 1420 1425 1430 1435	16983
ctc acc gag agg ctt tat gtt ggg ggc cct ctt acc aat tca agg ggg Leu Thr Glu Arg Leu Tyr Val Gly Gly Pro Leu Thr Asn Ser Arg Gly 1440 1445 1450	17031
gag aac tgc ggc tat cgc agg tgc cgc gcg agc ggc gta ctg aca act Glu Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr 1455 1460 1465	17079
agc tgt ggt aac acc ctc act tgc tac atc aag gcc cgg gca gcc tgt Ser Cys Gly Asn Thr Leu Thr Cys Tyr Ile Lys Ala Arg Ala Ala Cys 1470 1475 1480	17127
cga gcc gca ggg ctc cag gac tgc acc atg ctc gtg tgt ggc gac gac	17175

Arg Ala Ala Gly Leu Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp  
 1485 1490 1495

tta gtc gtt atc tgt gaa agc gcg ggg gtc cag gag gac gcg gcg agc 17223  
 Leu Val Val Ile Cys Glu Ser Ala Gly Val Gln Glu Asp Ala Ala Ser  
 1500 1505 1510 1515

ctg aga gcc ttc acg gag gct atg acc agg tac tcc gcc ccc cct ggg 17271  
 Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly  
 1520 1525 1530

gac ccc cca caa cca gaa tac gac ttg gag ctc ata aca tca tgc tcc 17319  
 Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser  
 1535 1540 1545

tcc aac gtg tca gtc gcc cac gac ggc gct gga aag agg gtc tac tac 17367  
 Ser Asn Val Ser Val Ala His Asp Gly Ala Gly Lys Arg Val Tyr Tyr  
 1550 1555 1560

ctc acc cgt gac cct aca acc ccc ctc gcg aga gct gcg tgg gag aca 17415  
 Leu Thr Arg Asp Pro Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr  
 1565 1570 1575

gca aga cac act cca gtc aat tcc tgg cta ggc aac ata atc atg ttt 17463  
 Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Phe  
 1580 1585 1590 1595

gcc ccc aca ctg tgg gcg agg atg ata ctg atg acc cat ttc ttt agc 17511  
 Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser  
 1600 1605 1610

gtc ctt ata gcc agg gac cag ctt gaa cag gcc ctc gat tgc gag atc 17559  
 Val Leu Ile Ala Arg Asp Gln Leu Glu Gln Ala Leu Asp Cys Glu Ile  
 1615 1620 1625

tac ggg gcc tgc tac tcc ata gaa cca ctg gat cta cct cca atc att 17607  
 Tyr Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro Pro Ile Ile  
 1630 1635 1640

caa aga ctc cat ggc ctc agc gca ttt tca ctc cac agt tac tct cca 17655  
 Gln Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro  
 1645 1650 1655

ggt gaa atc aat agg gtg gcc gca tgc ctc aga aaa ctt ggg gta ccg 17703  
 Gly Glu Ile Asn Arg Val Ala Ala Cys Leu Arg Lys Leu Gly Val Pro  
 1660 1665 1670 1675

ccc ttg cga gct tgg aga cac cgg gcc cgg agc gtc cgc gct agg ctt 17751  
 Pro Leu Arg Ala Trp Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu  
 1680 1685 1690

ctg gcc aga gga ggc agg gct gcc ata tgt ggc aag tac ctc ttc aac 17799  
 Leu Ala Arg Gly Gly Arg Ala Ala Ile Cys Gly Lys Tyr Leu Phe Asn  
 1695 1700 1705

tgg gca gta aga aca aag ctc aaa ctc act cca ata gcg gcc gct ggc 17847  
 Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Ala Ala Gly

DNA sequence

1710	1715	1720	
cag ctg gac ttg tcc ggc tgg ttc acg gct ggc tac agc ggg gga gac Gln Leu Asp Leu Ser Gly Trp Phe Thr Ala Gly Tyr Ser Gly Gly Asp			17895
1725	1730	1735	
att tat cac agc gtg tct cat gcc cgg ccc cgc tgg atc tgg ttt tgc Ile Tyr His Ser Val Ser His Ala Arg Pro Arg Trp Ile Trp Phe Cys			17943
1740	1745	1750	1755
cta ctc ctg ctt gct gca ggg gta ggc atc tac ctc ctc ccc aac cga Leu Leu Leu Ala Ala Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg			17991
1760	1765	1770	
atg agc acg aat cct aaa cct caa aga aag acc aaa cgt aac acc aac Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn			18039
1775	1780	1785	
cgg cgg ccg cag gac gtc aag ttc cgg ggt ggc ggt cag atc gtt ggt Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gln Ile Val Gly			18087
1790	1795	1800	
gga gtt tac ttg ttg ccg cgc agg ggc cct aga ttg ggt gtg cgc gcg Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala			18135
1805	1810	1815	
acg aga aag act tcc gag cgg tcg caa cct cga ggt aga cgt cag cct Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro			18183
1820	1825	1830	1835
atc ccc aag gct cgt cgg ccc gag ggc agg acc tgg gct cag ccc ggg Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly			18231
1840	1845	1850	
tac cct tgg ccc ctc tat ggc aat gag ggc tgc ggg tgg gcg gga tgg Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp			18279
1855	1860	1865	
ctc ctg tct ccc cgt ggc tct cgg cct agc tgg ggc ccc aca gac ccc Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro			18327
1870	1875	1880	
cgg cgt agg tcg cgc aat ttg ggt aag gtc atc gat acc ctt acg tgc Arg Arg Arg Ser Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys			18375
1885	1890	1895	
ggc ttc gcc gac ctc atg ggg tac ata ccg ctc gtc taatagtcga Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Leu Val			18421
1900	1905	1910	
cttggcccc actgtacttt tagctcgtaaaaatacat atactttca tttctccgta 18481			
aacaacatgt tttccatgt aatatccttt tctatTTTC gttccgttac caactttaca 18541			
catactttat atagctattc acttctatac actaaaaaaac taagacaatt ttaatTTGc 18601			
tgcctgccat attcaattt gttataaattt cctataattt atcctatttag tagctaaaa 18661			

0022472222222222

aagatgaatg tgaatcgaa cctaagagaa ttggatctga tccacaggac ggggtggc 18721  
gccatgatcg cgtagtcgat agtggctcca agtagcgaag cgagcaggac tggccggcgg 18781  
ccaaagcggt cggacagtgc tccgagaacg ggtgcgcata gaaattgcat caacgcata 18841  
agcgctagca gcacgccata gtgactggcg atgctgtcg aatggacgat atcccgaag 18901  
aggccccgca gtaccggcat aaccaagcct atgcctacag catccagggt gacggtgccg 18961  
aggatgacga tgagcgcatt gtttagattc atacacggtg cctgactgcg ttagcaattt 19021  
aactgtgata aactaccgca ttaaagcttt ttcttccaa tttttttttt ttgcgtcatta 19081  
taaaaatcat tacgaccgag attcccggt aataactgat ataattaaat tgaagctcta 19141  
atttgtgagt ttagtataca tgcatttact tataatacag ttttttagtt ttgcgtggccg 19201  
catcttctca aatatgcttc ccagcctgct tttctgtAAC gttcacccctc taccttagca 19261  
tcccttcctt ttgcaaatacg tcctttccaa acaataataa tgcagatcc tgcagagacc 19321  
acatcatcca cggttctata ctgttgaccc aatgcgtctc cccgtcata taaacccaca 19381  
ccgggtgtca taatcaacca atcgtaacct tcattcttc cacccatgac tctttgagca 19441  
ataaagccga taacaaaatc tttgtcgctc ttgcataatgt caacagtacc ctttagtat 19501  
tctccagtag ataggagcc ttgcgtac aattctgctt acatcaaaaag gcctcttagt 19561  
tcctttgtta cttcttctgc cgcctgctt aaaccgctaa caatacctgg gcccaccaca 19621  
ccgtgtgtcat tcgtaatgtc tgcccattct gctattctgt atacacccgc agagtactgc 19681  
aatttgactg tattaccaat gtcagcaaat ttctgtctt cgaagagtaa aaaattgtac 19741  
ttggcggata atgccttag cggcttaact gtgccttcca tggaaaaatc agtcaagata 19801  
tccacatgtt ttttagtaa acaaattttg ggacctaattt cttcaactaa ctccagtaat 19861  
tccttgggtgg tacgaacatc caatgaagca cacaagttt gttcttttc gtgcgtata 19921  
ttaaatagct tggcagcaac aggacttaga tgagtagcag cacgttcctt atatgtact 19981  
ttcgacatga ttatcttcg ttccctgcag gttttgttc tgtcagttt ggttaagaat 20041  
actgggcaat ttcatgtttc ttcaacacta catatgcgtt tatataccaa tctaagtctg 20101  
tgctccttcc ttctgttcc cttctgtcg gagattaccg aatcaaaaaaa atttcaagga 20161  
aacggaaatc aaaaaaaaaa ataaaaaaaaa aatgatgaat tgaaaagctt atcgat 20217

<210> 17  
<211> 1911  
<212> PRT  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:  
pd.delta.NS3NS5.pj.core140

<400> 17

Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro  
1 5 10 15

Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His  
20 25 30

Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly  
35 40 45

Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly  
50 55 60

Cys Ser Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser  
65 70 75 80

Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Ala  
85 90 95

Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala Thr Ala Thr Pro Pro  
100 105 110

Gly Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser  
115 120 125

Thr Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val  
130 135 140

Ile Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Lys Lys Cys  
145 150 155 160

Asp Glu Leu Ala Ala Lys Leu Val Ala Leu Gly Ile Asn Ala Val Ala  
165 170 175

Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val  
180 185 190

Val Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe  
195 200 205

Asp Ser Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe  
210 215 220

Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr Ile Thr Leu Pro Gln Asp  
225 230 235 240

Ala Val Ser Arg Thr Gln Arg Arg Gly Arg Thr Gly Arg Gly Lys Pro  
245 250 255

Gly Ile Tyr Arg Phe Val Ala Pro Gly Glu Arg Pro Ser Gly Met Phe  
260 265 270

Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp Ala Gly Cys Ala Trp Tyr  
275 280 285

Glu Leu Thr Pro Ala Glu Thr Thr Val Arg Leu Arg Ala Tyr Met Asn  
290 295 300

Thr Pro Gly Leu Pro Val Cys Gln Asp His Leu Glu Phe Trp Glu Gly  
305 310 315 320

Val Phe Thr Gly Leu Thr His Ile Asp Ala His Phe Leu Ser Gln Thr  
325 330 335

Lys Gln Ser Gly Glu Asn Leu Pro Tyr Leu Val Ala Tyr Gln Ala Thr  
340 345 350

Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser Trp Asp Gln Met Trp  
355 360 365

Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His Gly Pro Thr Pro Leu  
370 375 380

Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Ile Thr Leu Thr His Pro  
385 390 395 400

Val Thr Lys Tyr Ile Met Thr Cys Met Ser Ala Asp Leu Glu Val Val  
405 410 415

Thr Ser Thr Trp Val Leu Val Gly Gly Val Ile Ala Ala Leu Ala Ala  
420 425 430

Tyr Cys Leu Ser Thr Gly Cys Val Val Ile Val Gly Arg Val Val Leu  
435 440 445

Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu Val Leu Tyr Arg Glu  
450 455 460

Phe Asp Glu Met Glu Glu Cys Ser Gln His Leu Pro Tyr Ile Glu Gln  
465 470 475 480

Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys Ala Leu Gly Leu Leu  
485 490 495

Gln Thr Ala Ser Arg Gln Ala Glu Val Ile Ala Pro Ala Val Gln Thr  
500 505 510

Asn Trp Gln Lys Leu Glu Thr Phe Trp Ala Lys His Met Trp Asn Phe  
515 520 525

Ile Ser Gly Ile Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn  
530 535 540

Pro Ala Ile Ala Ser Leu Met Ala Phe Thr Ala Ala Val Thr Ser Pro  
545 550 555 560

Leu Thr Thr Ser Gln Thr Leu Leu Phe Asn Ile Leu Gly Gly Trp Val  
565 570 575

Ala Ala Gln Leu Ala Ala Pro Gly Ala Ala Thr Ala Phe Val Gly Ala  
 580 585 590

Gly Leu Ala Gly Ala Ala Ile Gly Ser Val Gly Leu Gly Lys Val Leu  
 595 600 605

Ile Asp Ile Leu Ala Gly Tyr Gly Ala Gly Val Ala Gly Ala Leu Val  
 610 615 620

Ala Phe Lys Ile Met Ser Gly Glu Val Pro Ser Thr Glu Asp Leu Val  
 625 630 635 640

Asn Leu Leu Pro Ala Ile Leu Ser Pro Gly Ala Leu Val Val Gly Val  
 645 650 655

Val Cys Ala Ala Ile Leu Arg Arg His Val Gly Pro Gly Glu Gly Ala  
 660 665 670

Val Gln Trp Met Asn Arg Leu Ile Ala Phe Ala Ser Arg Gly Asn His  
 675 680 685

Val Ser Pro Thr His Tyr Val Pro Glu Ser Asp Ala Ala Ala Arg Val  
 690 695 700

Thr Ala Ile Leu Ser Ser Leu Thr Val Thr Gln Leu Leu Arg Arg Leu  
 705 710 715 720

His Gln Trp Ile Ser Ser Glu Cys Thr Thr Pro Cys Ser Gly Ser Trp  
 725 730 735

Leu Arg Asp Ile Trp Asp Trp Ile Cys Glu Val Leu Ser Asp Phe Lys  
 740 745 750

Thr Trp Leu Lys Ala Lys Leu Met Pro Gln Leu Pro Gly Ile Pro Phe  
 755 760 765

Val Ser Cys Gln Arg Gly Tyr Lys Gly Val Trp Arg Gly Asp Gly Ile  
 770 775 780

Met His Thr Arg Cys His Cys Gly Ala Glu Ile Thr Gly His Val Lys  
 785 790 795 800

Asn Gly Thr Met Arg Ile Val Gly Pro Arg Thr Cys Arg Asn Met Trp  
 805 810 815

Ser Gly Thr Phe Pro Ile Asn Ala Tyr Thr Thr Gly Pro Cys Thr Pro  
 820 825 830

Leu Pro Ala Pro Asn Tyr Thr Phe Ala Leu Trp Arg Val Ser Ala Glu  
 835 840 845

Glu Tyr Val Glu Ile Arg Gln Val Gly Asp Phe His Tyr Val Thr Gly  
 850 855 860

Met Thr Thr Asp Asn Leu Lys Cys Pro Cys Gln Val Pro Ser Pro Glu  
 865 870 875 880

Phe Phe Thr Glu Leu Asp Gly Val Arg Leu His Arg Phe Ala Pro Pro  
885 890 895

Cys Lys Pro Leu Leu Arg Glu Glu Val Ser Phe Arg Val Gly Leu His  
900 905 910

Glu Tyr Pro Val Gly Ser Gln Leu Pro Cys Glu Pro Glu Pro Asp Val  
915 920 925

Ala Val Leu Thr Ser Met Leu Thr Asp Pro Ser His Ile Thr Ala Glu  
930 935 940

Ala Ala Gly Arg Arg Leu Ala Arg Gly Ser Pro Pro Ser Val Ala Ser  
945 950 955 960

Ser Ser Ala Ser Gln Leu Ser Ala Pro Ser Leu Lys Ala Thr Cys Thr  
965 970 975

Ala Asn His Asp Ser Pro Asp Ala Glu Leu Ile Glu Ala Asn Leu Leu  
980 985 990

Trp Arg Gln Glu Met Gly Gly Asn Ile Thr Arg Val Glu Ser Glu Asn  
995 1000 1005

Lys Val Val Ile Leu Asp Ser Phe Asp Pro Leu Val Ala Glu Glu Asp  
1010 1015 1020

Glu Arg Glu Ile Ser Val Pro Ala Glu Ile Leu Arg Lys Ser Arg Arg  
1025 1030 1035 1040

Phe Ala Gln Ala Leu Pro Val Trp Ala Arg Pro Asp Tyr Asn Pro Pro  
1045 1050 1055

Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr Glu Pro Pro Val Val His  
1060 1065 1070

Gly Cys Pro Leu Pro Pro Lys Ser Pro Pro Val Pro Pro Pro Arg  
1075 1080 1085

Lys Lys Arg Thr Val Val Leu Thr Glu Ser Thr Leu Ser Thr Ala Leu  
1090 1095 1100

Ala Glu Leu Ala Thr Arg Ser Phe Gly Ser Ser Ser Thr Ser Gly Ile  
1095 1110 1115 1120

Thr Gly Asp Asn Thr Thr Ser Ser Glu Pro Ala Pro Ser Gly Cys  
1125 1130 1135

Pro Pro Asp Ser Asp Ala Glu Ser Tyr Ser Ser Met Pro Pro Leu Glu  
1140 1145 1150

Gly Glu Pro Gly Asp Pro Asp Leu Ser Asp Gly Ser Trp Ser Thr Val  
1155 1160 1165

Ser Ser Glu Ala Asn Ala Glu Asp Val Val Cys Cys Ser Met Ser Tyr  
1170 1175 1180

Ser Trp Thr Gly Ala Leu Val Thr Pro Cys Ala Ala Glu Glu Gln Lys  
185 1190 1195 1200

Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu Arg His His Asn Leu  
1205 1210 1215

Val Tyr Ser Thr Thr Ser Arg Ser Ala Cys Gln Arg Gln Lys Lys Val  
1220 1225 1230

Thr Phe Asp Arg Leu Gln Val Leu Asp Ser His Tyr Gln Asp Val Leu  
1235 1240 1245

Lys Glu Val Lys Ala Ala Ser Lys Val Lys Ala Asn Leu Leu Ser  
1250 1255 1260

Val Glu Glu Ala Cys Ser Leu Thr Pro Pro His Ser Ala Lys Ser Lys  
265 1270 1275 1280

Phe Gly Tyr Gly Ala Lys Asp Val Arg Cys His Ala Arg Lys Ala Val  
1285 1290 1295

Thr His Ile Asn Ser Val Trp Lys Asp Leu Leu Glu Asp Asn Val Thr  
1300 1305 1310

Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu Val Phe Cys Val Gln  
1315 1320 1325

Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu Ile Val Phe Pro Asp  
1330 1335 1340

Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu Tyr Asp Val Val Thr  
345 1350 1355 1360

Lys Leu Pro Leu Ala Val Met Gly Ser Ser Tyr Gly Phe Gln Tyr Ser  
1365 1370 1375

Pro Gly Gln Arg Val Glu Phe Leu Val Gln Ala Trp Lys Ser Lys Lys  
1380 1385 1390

Thr Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe Asp Ser Thr Val  
1395 1400 1405

Thr Glu Ser Asp Ile Arg Thr Glu Glu Ala Ile Tyr Gln Cys Cys Asp  
1410 1415 1420

Leu Asp Pro Gln Ala Arg Val Ala Ile Lys Ser Leu Thr Glu Arg Leu  
425 1430 1435 1440

Tyr Val Gly Gly Pro Leu Thr Asn Ser Arg Gly Glu Asn Cys Gly Tyr  
1445 1450 1455

Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr Ser Cys Gly Asn Thr  
1460 1465 1470

Leu Thr Cys Tyr Ile Lys Ala Arg Ala Ala Cys Arg Ala Ala Gly Leu  
1475 1480 1485

Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp Leu Val Val Ile Cys  
1490 1495 1500

Glu Ser Ala Gly Val Gln Glu Asp Ala Ala Ser Leu Arg Ala Phe Thr  
505 1510 1515 1520

Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp Pro Pro Gln Pro  
1525 1530 1535

Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser Asn Val Ser Val  
1540 1545 1550

Ala His Asp Gly Ala Gly Lys Arg Val Tyr Tyr Leu Thr Arg Asp Pro  
1555 1560 1565

Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr Ala Arg His Thr Pro  
1570 1575 1580

Val Asn Ser Trp Leu Gly Asn Ile Ile Met Phe Ala Pro Thr Leu Trp  
585 1590 1595 1600

Ala Arg Met Ile Leu Met Thr His Phe Phe Ser Val Leu Ile Ala Arg  
1605 1610 1615

Asp Gln Leu Glu Gln Ala Leu Asp Cys Glu Ile Tyr Gly Ala Cys Tyr  
1620 1625 1630

Ser Ile Glu Pro Leu Asp Leu Pro Pro Ile Ile Gln Arg Leu His Gly  
1635 1640 1645

Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro Gly Glu Ile Asn Arg  
1650 1655 1660

Val Ala Ala Cys Leu Arg Lys Leu Gly Val Pro Pro Leu Arg Ala Trp  
665 1670 1675 1680

Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu Leu Ala Arg Gly Gly  
1685 1690 1695

Arg Ala Ala Ile Cys Gly Lys Tyr Leu Phe Asn Trp Ala Val Arg Thr  
1700 1705 1710

Lys Leu Lys Leu Thr Pro Ile Ala Ala Ala Gly Gln Leu Asp Leu Ser  
1715 1720 1725

Gly Trp Phe Thr Ala Gly Tyr Ser Gly Gly Asp Ile Tyr His Ser Val  
1730 1735 1740

Ser His Ala Arg Pro Arg Trp Ile Trp Phe Cys Leu Leu Leu Ala  
745 1750 1755 1760

Ala Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg Met Ser Thr Asn Pro  
1765 1770 1775

Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Gln Asp  
1780 1785 1790

Val Lys Phe Pro Gly Gly Gln Ile Val Gly Gly Val Tyr Leu Leu  
1795 1800 1805

Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala Thr Arg Lys Thr Ser  
1810 1815 1820

Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro Ile Pro Lys Ala Arg  
825 1830 1835 1840

Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly Tyr Pro Trp Pro Leu  
1845 1850 1855

Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp Leu Leu Ser Pro Arg  
1860 1865 1870

Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro Arg Arg Arg Ser Arg  
1875 1880 1885

Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu  
1890 1895 1900

Met Gly Tyr Ile Pro Leu Val  
905 1910

<210> 18  
<211> 20247  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
pd.delta.NS3NS5.pj.core150

<220>  
<221> CDS  
<222> (12679)..(18441)

<400> 18  
atcgatccta ccccttgcgc taaagaagta tatgtgccta ctaacgcttg tctttgtctc 60  
tgtcaactaaa cactggatta ttactcccag atacttattt tggactaatt taaatgattt 120  
cgatcaacg ttcttaatat cgctgaatct tccacaattt atgaaagtag ctaggaagag 180  
gaattggat aaagtttttt tttttgtaaa cctcgaagta tactcaaacg aatttagtat 240  
tttctcagtg atctcccaga tgcttcacc ctcacttaga agtgctttaa gcatttttt 300  
actgtggcta tttcccttat ctgcttcttc cgatgattcg aactgtaatt gcaaactact 360  
tacaatatca gtgatatcag attgatgttt ttgtccatag taaggaataa ttgtaaattc 420  
ccaaggcagga atcaatttct ttaatgaggc ttccagaatt gttgctttt gcgtcttgta 480  
tttaaactgg agtgatttat tgadaatatc gaaactcagc gaattgctta tgatagtatt 540

DAGMELLE

atagctcatg aatgtggctc tcttgattgc tgccgtta tgtgtaatca tccaaacataa 600  
ataggtagt tcagcagcac ataatgctat ttctcacct gaaggcttt caaacctttc 660  
cacaactga cgaacaagca ccttaggtgg tgccccat aatatatcaa attgtggcat 720  
gcttagcgcc gatcttgtgt gcaattgata tctagttca actactctat ttatcttgta 780  
tcttcagta ttcaaacacg ctaactcgaa aaactaaactt taattgtcct gtttgtctcg 840  
cgttcttcg aaaaatgcac cggccgcga ttatttgac tgcggaaaata attggtaactg 900  
cggtatctc atttcattt taaaaatgc accttgctg ctttccta atttttagac 960  
ggccgcagg ttcgtttgc ggtactatct tgtataaaaa agttgtttt acatgtgatc 1020  
tgcacagatt ttataatgta ataagcaaga atacattatc aaacgaacaa tactggtaaa 1080  
agaaaaccaa aatggacgac attgaaacag ccaagaatct gacggtaaaa gcacgtacag 1140  
cttatagcgt ctgggatgta tgcggctgt ttattgaaat gattgtcct gatgtgatc 1200  
ttgatataaga gagtaaacgt aagtctgatg agctactctt tccaggatat gtcataaggc 1260  
ccatggaatc ttcacaacc ggtaggccgt atggcttga ttctagcgca gaagattcca 1320  
gcgtatctc tgactccagt gctgaggtaa tttgcctgc tgcgaagatg gttaaggaaa 1380  
ggtttattt gattggaaat ggtatgctt cttcacaaga agcaagtcg gctgccatag 1440  
atttgcgt acagaataac aagctgttag acaatagaaa gcaactatac aaatctattt 1500  
ctataataat aggaagattt cccgagaaag acaagaagag agctaccgaa atgctcatga 1560  
aaaaaatgga ttgtacacag ttattgtcc caccagttcc aacggaagaa gatgttatga 1620  
agctcgtaag cgctgttacc caattgctt ctttagttcc accagatcgt caagctgctt 1680  
taataggta ttatttcattt ccggaaatctc taaaggatat attcaatagt ttcaatgaac 1740  
tggccggcaga gaatcgaaa cagaaaaaaa agatgtgat ggaaggaagg actgaagtga 1800  
accatgctaa tacaaatgaa gaagttccct ccaggcgaac aagaagttaga gacacaaatg 1860  
caagaggagc atataaatta caaaacacca tcactgaggg ccctaaagcg gttcccacga 1920  
aaaaaaggag agtagcaacg aggtaaggg gcagaaaatc acgtaataact tctagggtat 1980  
gatccaatat caaaggaaat gatagcatg aaggatgaga ctaatccat tgaggagtgg 2040  
cagcatatac aacagctaaa ggtagtgct gaaggaagca tacgataccc cgcatgaaat 2100  
gggataatat cacaggaggt actagactac ctttcatttcc acataaatacg acgcatataa 2160  
gtacgcattt aagcataaac acgactatg ccgttcttct catgtatata tatatacagg 2220  
caacacgcag atataaggta gacgtgaaca gtgagctgta tgtgcgcagc tcgcgttgca 2280

0102030405060708090

tttcggaag cgctcgaaa cgaaaaacgct ttgaagttcc tattccgaaat ttcctattct 2340  
ctagaaaatgt taggaacttc agagcgcttt tgaaaaccaa aagcgctctg aagacgcact 2400  
ttcaaaaaaac caaaaacgca ccggactgta acgagctact aaaatattgc gaataccgct 2460  
tccacaaaca ttgctcaaaa gatatcttt gctatatac tctgtgctat atccctatat 2520  
aacctaccca tccacccccc gtccttgaa cttgcataa aactcgacct ctacatcaac 2580  
aggcttccaa tgctctcaa atttactgt caagtagacc catacggtg taatatgctg 2640  
ctcttcataa tgtaagctta tctttatcga atcgtgtgaa aaactactac cgcgataaac 2700  
ctttacgggtt ccctgagatt gaatttagttc cttagtata tgataacaaga cactttgaa 2760  
ctttagtacga cgaattttga ggttcgccat cctctggcta tttccatta tcctgtcggc 2820  
tattatctcc gcctcagttt gatctccgc ttcagactgc cattttcac ataatgaatc 2880  
tatttcaccc cacaatcctt catccgcctc cgcatcttgt tccgttaaac tattgacttc 2940  
atgttgtaca ttgttagtt cacgagaagg gtcctttca ggcggtagct cctgatctcc 3000  
tatatgacct ttatcctgtt ctcttccac aaacttagaa atgtattcat gaattatgga 3060  
gcacctaata acattttca aggccggagaa gtttggccaa gatgccaaat atgcttgaca 3120  
tgaaaacgtg agaatgaatt tagtattatt gtgatattct gaggcaattt tattataatc 3180  
tcgaagataa gagaagaatg cagtgaccc ttgtattgaca aatggagatt ccatgtatct 3240  
aaaaaaatacg cctttaggcc ttctgatacc ctccccctg cggttagctg tgcctttac 3300  
attaatatct aaaccctctc cgatggtggc cttaactga ctaataatg caaccgatat 3360  
aaactgtgat aattctgggt gatttatgat tcgatcgaca attgtattgt acactagtgc 3420  
aggatcagggc caatccagtt ctccccat taccgggtgtg tcgtctgtat tcagtgatcg 3480  
tccaacaaat gcaaattgcta acgtttgtt tttcttataa ttgtcaggaa ctggaaaatg 3540  
cccccttgcgtc gtctcgatata cacacctact ttcatcgat accatagggtt ggaagtgcgt 3600  
cataatacat tgcttaatac aagcaaggcag tctctcgcca ttcatatttc agttatccc 3660  
cattacagct gatgtcattt gatattcagcg ctgtaaaaat ctatctgtt cagaagggtt 3720  
tcgcggtttt tataaacaaa actttcgta cgaaatcgag caatcaccctt agctgcgtat 3780  
ttggaaaattc gggaaaaatg agagcaacgc gagttgcatt ttttacacca taatgcata 3840  
ttaacttcga gaagggtttaa aggcttaattt cactgtatg tttcaaaaaac ctcaatctgt 3900  
ccattgaatg ccttataaaa cagctataga ttgcataagaa gagttagctt ctcaatgctt 3960  
tttgcataaag cttactgtatg atgatgtgtc tactttcagg cgggtctgtt gtaaggagaa 4020

tgacattata aagctggcac tttagaattcc acggactata gactatacta gtatactccg 4080  
tctactgtac gatacacttc cgctcaggc cttgcctt aacgaggc taccactctt 4140  
ttgttactct attgatccag ctcagcaaag gcagtgtat ctaagattct atttcgca 4200  
tgtagtaaaa ctagctagac cgagaaagag actagaaatg caaaaggcac ttctacaatg 4260  
gctgccatca ttattatccg atgtgacgct gcatttttt tttttttt tttttttt 4320  
ttttttttt tttttttt tttttggta caaatatcat aaaaaaagag aatctttta 4380  
agcaaggatt ttcttaactt ctgcggcac agcatcaccg acttcggtgg tactgttgaa 4440  
accacctaata tcaccagttc tgataacctgc atccaaaacc ttttaactg catcttcaat 4500  
ggcttaccc tcttcaggca agttcaatga caattcaac atcattgcag cagacaagat 4560  
agtggcgata gggttgaccc tattcttgg caaatctggc gcggaaaccat ggcattggc 4620  
gtacaaacca aatgcggtgt tcttgcgttgg caaagaggcc aaggacgcag atggcaacaa 4680  
acccaaggag cctggataa cggaggcttc atcggagatg atatcaccaa acatgttgct 4740  
ggtgattata ataccattta ggtgggttgg gttcttaact aggatcatgg cggcagaatc 4800  
aatcaattga tggtgaactt tcaatgttagg gaattcggtt ttgtatggg cctccacagt 4860  
tttctccat aatcttgaag aggccaaaac attagctta tccaaggacc aaataggcaa 4920  
tggggctca tggttaggg ccatgaaagc ggccattttt gtgattctt gcacttctgg 4980  
aacgggttat tggtcactat cccaaagcgac accatcacca tcgtcttcc ttcttacc 5040  
aaagtaaata cctccacta attctctaac aacaacgaag tcagttaccc tagcaaattg 5100  
tggcttgatt ggagataagt ctaaaagaga gtcggatgca aagttacatg gtcttaagtt 5160  
ggcgtacaat tgaagttctt tacggatttt tagtaaacct tggtcaggc taacactacc 5220  
ggtaccccat ttaggaccac ccacagcacc taacaaaacg gcatcaggc tcttggaggc 5280  
ttccagcgcc tcatctggaa gtggAACACC tgtagcatcg atagcagcac caccaattaa 5340  
atgattttcg aaatcgaact tgacattggc acggacatca gaaatagctt taagaacctt 5400  
aatggcttcg gctgtgattt cttgaccaac gtggcacct ggcaaaacga cgatcttctt 5460  
aggggcagac attacaatgg tatatccttg aaatatatat aaaaaaaaaa aaaaaaaaaa 5520  
aaaaaaaaaa atgcagcttc tcaatgatat tcgaatacgc tttgaggaga tacagcctaa 5580  
tatccgacaa actgtttac agatttacga tcgtacttgt tacccatcat tgaattttga 5640  
acatccgaac ctgggagttt tccctgaaac agatagtata tttgaacctg tataataata 5700  
tatagtctag cgcttacgg aagacaatgt atgtatccg gttcctggag aaactattgc 5760



attractcgt ttttaggactg gttcagaatt cctcaaaaat tcatccaaat atacaagtgg 7560  
atcgatgata agctgtcaaa catgagaatt cttgaagacg aaagggcctc gtgatacgcc 7620  
tattttata ggttaatgtc atgataataa tggtttctt gacgtcaggt ggcactttc 7680  
ggggaaatgt gcgcgaaacc cctatttgtt tattttcta aatacattca aatatgtatc 7740  
cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg aagagtatga 7800  
gtattcaaca tttccgtgtc gcccttattc ccttttgc ggcatttgc cttcctgttt 7860  
ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttt ggtgcacgag 7920  
tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagttt cgccccgaag 7980  
aacgtttcc aatgatgagc acttttaaag ttctgctatg tggcgcggtt ttatccgtg 8040  
ttgacgcccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggtt 8100  
agtactcacc agtcacagaa aagcatctt cggatggcat gacagtaaga gaattatgca 8160  
gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca acgatcgag 8220  
gaccgaagga gctaaccgct ttttgcaca acatggggga tcatgtact cgccttgatc 8280  
gttgggaacc ggagctgaat gaagccatac caaacgacga cgcgtgacacc acgatgcctg 8340  
cagcaatggc aacaacgttg cgcaaactat taactggcga actacttact ctagttccc 8400  
ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg 8460  
ccctccggc tggctggttt attgctgata aatctggagc cggtgagcgt gggtctcgcg 8520  
gtatcattgc agcaactggg ccagatggta agccctcccg tatcgttagtt atctacacga 8580  
cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac 8640  
tgattaagca ttggtaactg tcagaccaag tttactcata tatactttttag attgatttaa 8700  
aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgcataat ctcatgacca 8760  
aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag 8820  
gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaaccac 8880  
cgctaccagc ggtgggttgt ttgcggatc aagagctacc aactctttt ccgaaggtaa 8940  
ctggcttcag cagagcgcag ataccaaata ctgtccttct agttagccg tagttaggcc 9000  
accacttcaa gaactctgta gcaccgccta cataacctgc tctgctaatac ctgttaccag 9060  
tggctgctgc cagtggcgt aagtgcgtgtc ttaccgggtt ggactcaaga cgatagttac 9120  
cggataaggc gcagcggcgtc ggctgaacgg ggggttcgtg cacacagccc agcttggagc 9180  
gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc 9240

DNA sequence analysis

ccgaaggagaaaggcggac aggtatccgg taagcggcag ggtcggaaca ggagagcgca 9300  
cgagggagct tccaggggaa aacgccttgtt atctttatag tcctgtcggg tttcgccacc 9360  
tctgacttga gcgtcgattt ttgtgatgt cgtcaggggg gcgagccta tggaaaaacg 9420  
ccagcaacgc ggcctttta cggttcctgg cctttgctg gcctttgct cacatgttct 9480  
ttcctgcgtt atcccgtat tctgtggata accgttattac cgccttgag tgagctgata 9540  
ccgctcgccg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa gcggaagagc 9600  
gcctgatgcg gtatttctc cttacgcattc tgtcggtat ttcacaccgc atatggtgc 9660  
ctctcagtac aatctgctct gatgccgcat agttaagcca gtatacaccc cgctatcgct 9720  
acgtgactgg gtcatggctg cgccccgaca cccgccaaca cccgctgacg cgccctgacg 9780  
ggcttgcgtc ctccggcat ccgcttacag acaagctgtg accgtctccg ggagctgcat 9840  
gtgtcagagg ttttaccgt catcaccgaa acgcgcgagg cagctcggt aaagctcatc 9900  
agcgtggtcg tgaagcgatt cacagatgtc tgcctgttca tccgcgtcca gctcggttag 9960  
tttctccaga agcgttaatg tctggcttct gataaagcgg ccgtttaa gggcggtttt 10020  
ttcctgtttt gtcactgtatg cctccgtgtt aaaaaaaaaaaaaat 10080  
accgatgaaa cgagagagga tgctcacat acgggtttact gatgtgaac atgcccgtt 10140  
actggAACGT tgtgagggtt aacaactggc ggtatggatg cggcgggacc agagaaaaat 10200  
cactcagggt caatgccagc gcttcgttaa tacagatgtt ggtgttccac agggtagcca 10260  
gcagcatcct gcgtgcaga tccggaacat aatggtgcag ggcgttact tccgcgttcc 10320  
cagactttac gaaacacgga aaccgaagac cattcatgtt gttgctcagg tcgcagacgt 10380  
tttgcagcag cagtcgttcc acgttcgttc ggtatcggt gattcattct gctaaccagt 10440  
aaggcaaccc cgccagccta gccgggtcct aacgacagg agcacgtatca tgccaccccg 10500  
tggccaggac ccaacgctgc ccgagatgcg ccgcgtgcgg ctgctggaga tggcggacgc 10560  
gatggatatg ttctgccaag gttggtttg cgcattcaca gttctccgca agaattgatt 10620  
ggctccaatt cttggagtgg tgaatccgtt agcgaggtgc cgccggcttc cattcaggc 10680  
gaggtggccc ggctccatgc accgcgacgc aacgcggggaa ggcagacaag gtatagggcg 10740  
gcgcctacaa tccatgccaatccat gtgctcgccg aggccgcata aatcgccgtg 10800  
acgatcagcg gtccaatgtat cgaagttagg ctggtaagag ccgcgagcga tccttgaagc 10860  
tgtccctgtat ggtcgcatc tacctgcctg gacagcatgg cctgcaacgc gggcatcccg 10920  
atgcggccgg aagcgagaag atcataatg gggaaaggcca tccagcctcg cgtcgcgaac 10980

gccagcaaga cgtagcccg agcgtcgccc gccatgcgg cgataatggc ctgcttcg 11040  
ccgaaacgtt tgggtggcggg accagtgacg aaggcttgag cgagggcgtg caagattccg 11100  
aataccgcaa gcgcacaggcc gatcatcgac gcgctccagc gaaagcggtc ctcggccaaa 11160  
atgacccaga gcgctgccgg cacctgtcct acgagttgca tgataaagaa gacagtcata 11220  
agtgcggcga cgatagtcat gccccgcgcc caccggaaagg agctgactgg gttgaaggct 11280  
ctcaaggca tcggtcgagg atccttcaat atgcgcacat acgctgttat gttcaaggctc 11340  
ccttcgttta agaacgaaag cggcttcct tttgagggat gttcaagtt gttcaaatct 11400  
atcaaatttg caaatccccca gtctgtatct agagcgttga atcggtgatg cgatttgtta 11460  
attaaattga tggtgtcacc attaccaggc ctagatatac caatggcaaa ctgagcacaa 11520  
caataccagt ccggatcaac tggcaccatc tctcccttag tctcatctaa ttttcttcc 11580  
ggatgagggtt ccagatatac cgcaacacct ttattatggt ttccctgagg gaataataga 11640  
atgtcccatt cgaaatcacc aattctaaac ctggcgaat tgtatccgg gtttgttaac 11700  
tcgttccagt caggaatgtt ccacgtgaag ctatcttcca gcaaagtctc cacttcttca 11760  
tcaaattgtg gagaatactc ccaatgtct tatctatggg acctccggga aacacagtac 11820  
cgatacttcc caattcgtct tcagagctca ttgtttgtt gaagagacta atcaaagaat 11880  
cgtttctca aaaaaattaa tatcttaact gatagtttga tcaaaggggc aaaacgtagg 11940  
ggcaaacaacaa cgaaaaatc gtttctcaa ttttctgatg ccaagaactc taaccagtct 12000  
tatctaaaaa ttgccttag atccgtctt ccgttacag cctgtgtaac tgattaatcc 12060  
tgccttctta atcaccattc taatgtttta attaaggat tttgtcttca ttaacggctt 12120  
tcgctcataa aaatgttag acgtttgcc cgccggcggg aaaccatcca cttcacgaga 12180  
ctgatctcct ctgcccgaac accgggcatc tccaaactt aagttggaga aataagagaa 12240  
tttcagattt agagaatgaa aaaaaaaaaac cttagttca taggtccatt ctcttagcgc 12300  
aactacagag aacaggggca caaacaggca aaaaacgggc acaacctcaa tggagtgtatg 12360  
caacctgcct ggagtaaatg atgacacaag gcaattgacc cacgtatgtt tctatctcat 12420  
tttcttacac cttcttattac cttctgtct ctctgatttg gaaaaagctg aaaaaaaaaagg 12480  
ttgaaaccag ttccctgaaa ttattccct acttgactaa taagtatata aagacggtag 12540  
gtattgattt taattctgtt aatcttatttc ttaaaacttct taaattctac ttttatagtt 12600  
agtctttttt ttagtttaa aacaccaaga acttagttc gaataaacac acataaacaa 12660  
acaagcttac aaaacaaa atg gct gca tat gca gct cag ggc tat aag gtg 12711

Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val			
1	5	10	
cta gta ctc aac ccc tct gtt gct gca aca ctg ggc ttt ggt gct tac			12759
Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr			
15	20	25	
atg tcc aag gct cat ggg atc gat cct aac atc agg acc ggg gtg aga			12807
Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg			
30	35	40	
aca att acc act ggc agc ccc atc acg tac tcc acc tac ggc aag ttc			12855
Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe			
45	50	55	
ctt gcc gac ggc ggg tgc tcg ggg ggc gct tat gac ata ata att tgt			12903
Leu Ala Asp Gly Gly Cys Ser Gly Ala Tyr Asp Ile Ile Ile Cys			
60	65	70	75
gac gag tgc cac tcc acg gat gcc aca tcc atc ttg ggc att ggc act			12951
Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr			
80	85	90	
gtc ctt gac caa gca gag act gcg ggg gcg aga ctg gtt gtg ctc gcc			12999
Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala			
95	100	105	
acc gcc acc cct ccg ggc tcc gtc act gtg ccc cat ccc aac atc gag			13047
Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn Ile Glu			
110	115	120	
gag gtt gct ctg tcc acc acc gga gag atc cct tt tac ggc aag gct			13095
Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala			
125	130	135	
atc ccc ctc gaa gta atc aag ggg ggg aga cat ctc atc ttc tgt cat			13143
Ile Pro Leu Glu Val Ile Lys Gly Arg His Leu Ile Phe Cys His			
140	145	150	155
tca aag aag aag tgc gac gaa ctc gcc gca aag ctg gtc gca ttg ggc			13191
Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Ala Leu Gly			
160	165	170	
atc aat gcc gtg gcc tac tac cgc ggt ctt gac gtg tcc gtc atc ccg			13239
Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile Pro			
175	180	185	
acc agc ggc gat gtt gtc gtc gtg gca acc gat gcc ctc atg acc ggc			13287
Thr Ser Gly Asp Val Val Val Ala Thr Asp Ala Leu Met Thr Gly			
190	195	200	
tat acc ggc gac ttc gac tgc gtg ata gac tgc aat acg tgt gtc acc			13335
Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys Val Thr			
205	210	215	
cag aca gtc gat ttc agc ctt gac cct acc ttc acc att gag aca atc			13383
Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr Ile			

DRAFT

220	225	230	235		
acg ctc ccc caa gat gct gtc tcc cgc act caa cgt cgg ggc agg act				13431	
Thr Leu Pro Gln Asp Ala Val Ser Arg Thr Gln Arg Arg Gly Arg Thr					
240	245	250			
ggc agg ggg aag cca ggc atc tac aga ttt gtg gca cgg ggg gag cgc				13479	
Gly Arg Gly Lys Pro Gly Ile Tyr Arg Phe Val Ala Pro Gly Glu Arg					
255	260	265			
ccc tcc ggc atg ttc gac tcg tcc gtc ctc tgt gag tgc tat gac gca				13527	
Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp Ala					
270	275	280			
ggc tgt gct tgg tat gag ctc acg ccc gcc gag act aca gtt agg cta				13575	
Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Thr Val Arg Leu					
285	290	295			
cga gcg tac atg aac acc ccg ggg ctt ccc gtg tgc cag gac cat ctt				13623	
Arg Ala Tyr Met Asn Thr Pro Gly Leu Pro Val Cys Gln Asp His Leu					
300	305	310	315		
gaa ttt tgg gag ggc gtc ttt aca ggc ctc act cat ata gat gcc cac				13671	
Glu Phe Trp Glu Gly Val Phe Thr Gly Leu Thr His Ile Asp Ala His					
320	325	330			
ttt cta tcc cag aca aag cag agt ggg gag aac ctt cct tac ctg gta				13719	
Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Leu Pro Tyr Leu Val					
335	340	345			
gcg tac caa gcc acc gtg tgc gct agg gct caa gcc cct ccc cca tcg				13767	
Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser					
350	355	360			
tgg gac cag atg tgg aag tgt ttg att cgc ctc aag ccc acc ctc cat				13815	
Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His					
365	370	375			
ggg cca aca ccc ctg cta tac aga ctg ggc gct gtt cag aat gaa atc				13863	
Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Ile					
380	385	390	395		
acc ctg acg cac cca gtc acc aaa tac atc atg aca tgc atg tcg gcc				13911	
Thr Leu Thr His Pro Val Thr Lys Tyr Ile Met Thr Cys Met Ser Ala					
400	405	410			
gac ctg gag gtc acg agc acc tgg gtg ctc gtt ggc ggc gtc ctg				13959	
Asp Leu Glu Val Val Thr Ser Thr Trp Val Leu Val Gly Gly Val Leu					
415	420	425			
gct gct ttg gcc gcg tat tgc ctg tca aca ggc tgc gtg gtc ata gtg				14007	
Ala Ala Leu Ala Ala Tyr Cys Leu Ser Thr Gly Cys Val Val Ile Val					
430	435	440			
ggc agg gtc gtc ttg tcc ggg aag ccg gca atc ata cct gac agg gaa				14055	
Gly Arg Val Val Leu Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu					
445	450	455			

DRAFT

gtc ctc tac cga gag ttc gat gag atg gaa gag tgc tct cag cac tta	14103
Val Leu Tyr Arg Glu Phe Asp Glu Met Glu Glu Cys Ser Gln His Leu	
460 465 470 475	
ccg tac atc gag caa ggg atg atg ctc gcc gag cag ttc aag cag aag	14151
Pro Tyr Ile Glu Gln Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys	
480 485 490	
gcc ctc ggc ctc ctg cag acc gcg tcc cgt cag gca gag gtt atc gcc	14199
Ala Leu Gly Leu Leu Gln Thr Ala Ser Arg Gln Ala Glu Val Ile Ala	
495 500 505	
cct gct gtc cag acc aac tgg caa aaa ctc gag acc ttc tgg gcg aag	14247
Pro Ala Val Gln Thr Asn Trp Gln Lys Leu Glu Thr Phe Trp Ala Lys	
510 515 520	
cat atg tgg aac ttc atc agt ggg ata caa tac ttg gcg ggc ttg tca	14295
His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala Gly Leu Ser	
525 530 535	
acg ctg cct ggt aac ccc gcc att gct tca ttg atg gct ttt aca gct	14343
Thr Leu Pro Gly Asn Pro Ala Ile Ala Ser Leu Met Ala Phe Thr Ala	
540 545 550 555	
gct gtc acc agc cca cta acc act agc caa acc ctc ctc ttc aac ata	14391
Ala Val Thr Ser Pro Leu Thr Ser Gln Thr Leu Leu Phe Asn Ile	
560 565 570	
ttg ggg ggg tgg gtg gct gcc cag ctc gcc gcc ccc ggt gcc gct act	14439
Leu Gly Gly Trp Val Ala Ala Gln Leu Ala Ala Pro Gly Ala Ala Thr	
575 580 585	
gcc ttt gtg ggc gct ggc tta gct ggc gcc atc ggc agt gtt gga	14487
Ala Phe Val Gly Ala Gly Leu Ala Gly Ala Ala Ile Gly Ser Val Gly	
590 595 600	
ctg ggg aag gtc ctc ata gac atc ctt gca ggg tat ggc gcg ggc gtg	14535
Leu Gly Lys Val Leu Ile Asp Ile Leu Ala Gly Tyr Gly Ala Gly Val	
605 610 615	
gcg gga gct ctt gtg gca ttc aag atc atg agc ggt gag gtc ccc tcc	14583
Ala Gly Ala Leu Val Ala Phe Lys Ile Met Ser Gly Glu Val Pro Ser	
620 625 630 635	
acg gag gac ctg gtc aat cta ctc ccc gcc atc ctc tcg ccc gga gcc	14631
Thr Glu Asp Leu Val Asn Leu Leu Pro Ala Ile Leu Ser Pro Gly Ala	
640 645 650	
ctc gta gtc ggc gtg gtc tgt gca gca ata ctg cgc cggt cac gtt ggc	14679
Leu Val Val Gly Val Val Cys Ala Ala Ile Leu Arg Arg His Val Gly	
655 660 665	
ccg ggc gag ggg gca gtg dag tgg atg aac cgg ctg ata gcc ttc gcc	14727
Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile Ala Phe Ala	
670 675 680	
tcc cgg ggg aac cat gtc tcc ccc acg cac tac gtg ccg gag agc gat	14775

Ser Arg Gly Asn His Val Ser Pro Thr His Tyr Val Pro Glu Ser Asp  
 685 690 695

gca gct gcc cgc gtc act gcc ata ctc agc agc ctc act gta acc cag 14823  
 Ala Ala Ala Arg Val Thr Ala Ile Leu Ser Ser Leu Thr Val Thr Gln  
 700 705 710 715

ctc ctg agg cga ctg cac cag tgg ata agc tcg gag tgt acc act cca 14871  
 Leu Leu Arg Arg Leu His Gln Trp Ile Ser Ser Glu Cys Thr Thr Pro  
 720 725 730

tgc tcc ggt tcc tgg cta agg gac atc tgg gac tgg ata tgc gag gtg 14919  
 Cys Ser Gly Ser Trp Leu Arg Asp Ile Trp Asp Trp Ile Cys Glu Val  
 735 740 745

ttg agc gac ttt aag acc tgg cta aaa gct aag ctc atg cca cag ctg 14967  
 Leu Ser Asp Phe Lys Thr Trp Leu Lys Ala Lys Leu Met Pro Gln Leu  
 750 755 760

cct ggg atc ccc ttt gtg tcc tgc cag cgc ggg tat aag ggg gtc tgg 15015  
 Pro Gly Ile Pro Phe Val Ser Cys Gln Arg Gly Tyr Lys Gly Val Trp  
 765 770 775

cga ggg gac ggc atc atg cac act cgc tgc cac tgt gga gct gag atc 15063  
 Arg Gly Asp Gly Ile Met His Thr Arg Cys His Cys Gly Ala Glu Ile  
 780 785 790 795

act gga cat gtc aaa aac ggg acg atg agg atc gtc ggt cct agg acc 15111  
 Thr Gly His Val Lys Asn Gly Thr Met Arg Ile Val Gly Pro Arg Thr  
 800 805 810

tgc agg aac atg tgg agt ggg acc ttc ccc att aat gcc tac acc acg 15159  
 Cys Arg Asn Met Trp Ser Gly Thr Phe Pro Ile Asn Ala Tyr Thr Thr  
 815 820 825

ggc ccc tgt acc ccc ctt cct gcg ccg aac tac acg ttc gcg cta tgg 15207  
 Gly Pro Cys Thr Pro Leu Pro Ala Pro Asn Tyr Thr Phe Ala Leu Trp  
 830 835 840

agg gtg tct gca gag gaa tac gtg gag ata agg cag gtg ggg gac ttc 15255  
 Arg Val Ser Ala Glu Glu Tyr Val Glu Ile Arg Gln Val Gly Asp Phe  
 845 850 855

cac tac gtg acg ggt atg act act gac aat ctt aaa tgc ccg tgc cag 15303  
 His Tyr Val Thr Gly Met Thr Asp Asn Leu Lys Cys Pro Cys Gln  
 860 865 870 875

gtc cca tcg ccc gaa ttt ttc aca gaa ttg gac ggg gtg cgc cta cat 15351  
 Val Pro Ser Pro Glu Phe Phe Thr Glu Leu Asp Gly Val Arg Leu His  
 880 885 890

agg ttt gcg ccc ccc tgc aag ccc ttg ctg cggt gag gag gta tca ttc 15399  
 Arg Phe Ala Pro Pro Cys Lys Pro Leu Leu Arg Glu Glu Val Ser Phe  
 895 900 905

aga gta gga ctc cac gaa tac ccg gta ggg tcg caa tta cct tgc gag 15447  
 Arg Val Gly Leu His Glu Tyr Pro Val Gly Ser Gln Leu Pro Cys Glu

910                    915                    920

ccc gaa ccg gac gtg gcc gtg ttg acg tcc atg ctc act gat ccc tcc Pro Glu Pro Asp Val Ala Val Leu Thr Ser Met Leu Thr Asp Pro Ser			15495
925                    930                    935			
cat ata aca gca gag gcg gcc ggg cga agg ttg gcg agg gga tca ccc His Ile Thr Ala Glu Ala Ala Gly Arg Arg Leu Ala Arg Gly Ser Pro			15543
940                    945                    950                    955			
ccc tct gtg gcc agc tcc tcg gct agc cag cta tcc gct cca tct ctc Pro Ser Val Ala Ser Ser Ala Ser Gln Leu Ser Ala Pro Ser Leu			15591
960                    965                    970			
aag gca act tgc acc gct aac cat gac tcc cct gat gct gag ctc ata Lys Ala Thr Cys Thr Ala Asn His Asp Ser Pro Asp Ala Glu Leu Ile			15639
975                    980                    985			
gag gcc aac ctc cta tgg agg cag gag atg ggc ggc aac atc acc agg Glu Ala Asn Leu Leu Trp Arg Gln Glu Met Gly Gly Asn Ile Thr Arg			15687
990                    995                    1000			
gtt gag tca gaa aac aaa gtg gtg att ctg gac tcc ttc gat ccg ctt Val Glu Ser Glu Asn Lys Val Val Ile Leu Asp/Ser Phe Asp Pro Leu			15735
1005                    1010                    1015			
gtg gcg gag gag gac gag cgg gag atc tcc gca ccc gca gaa atc ctg Val Ala Glu Glu Asp Glu Arg Glu Ile Ser Val Pro Ala Glu Ile Leu			15783
1020                    1025                    1030                    1035			
cgg aag tct cgg aga ttc gcc cag gcc ctg ccc gtt tgg gcg cgg cgg Arg Lys Ser Arg Arg Phe Ala Gln Ala Leu Pro Val Trp Ala Arg Pro			15831
1040                    1045                    1050			
gac tat aac ccc ccg cta gtg gag acg tgg aaa aag ccc gac tac gaa Asp Tyr Asn Pro Pro Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr Glu			15879
1055                    1060                    1065			
cca cct gtg gtc cat ggc tgc ccg ctt cca cct cca aag tcc cct cct Pro Pro Val Val His Gly Cys Pro Leu Pro Pro Pro Lys Ser Pro Pro			15927
1070                    1075                    1080			
gtg cct ccg cct cgg aag aag cgg acg gtg gtc ctc act gaa tca acc Val Pro Pro Pro Arg Lys Lys Arg Thr Val Val Leu Thr Glu Ser Thr			15975
1085                    1090                    1095			
cta tct act gcc ttg gcc gag ctc gcc acc aga acg ttt ggc agc tcc Leu Ser Thr Ala Leu Ala Glu Leu Ala Thr Arg Ser Phe Gly Ser Ser			16023
1100                    1105                    1110                    1115			
tca act tcc ggc att acg ggc gac aat acg aca aca tcc tct gag ccc Ser Thr Ser Gly Ile Thr Gly Asp Asn Thr Thr Ser Ser Glu Pro			16071
1120                    1125                    1130			
gcc cct tct ggc tgc ccc ccc gac tcc gac gct gag tcc tat tcc tcc Ala Pro Ser Gly Cys Pro Pro Asp Ser Asp Ala Glu Ser Tyr Ser Ser			16119
1135                    1140                    1145			

16167

atg ccc ccc ctg gag ggg gag cct ggg gat ccg gat ctt agc gac ggg  
 Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu Ser Asp Gly  
 1150 1155 1160

tca tgg tca acg gtc agt agt gag gcc aac gcg qag gat gtc gtg tgc 16215  
 Ser Trp Ser Thr Val Ser Ser Glu Ala Asn Ala Glu Asp Val Val Cys  
 1165 1170 1175

tgc tca atg tct tac tct tgg aca ggc gca ctc gtc acc ccg tgc gcc 16263  
 Cys Ser Met Ser Tyr Ser Trp Thr Gly Ala Leu Val Thr Pro Cys Ala  
 1180 1185 1190 1195

gcg gaa gaa cag aaa ctg ccc atc aat gca cta agc aac tgc ttg cta 16311  
 Ala Glu Glu Gln Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu  
 1200 1205 1210

cgt cac cac aat ttg gtg tat tcc acc acc tca cgc agt gct tgc caa 16359  
 Arg His His Asn Leu Val Tyr Ser Thr Ser Arg Ser Ala Cys Gln  
 1215 1220 1225

agg cag aag aaa gtc aca ttt gac aga ctg caa gtt ctg gac agc cat 16407  
 Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Ser His  
 1230 1235 1240

tac cag gac gta ctc aag gag gtt aaa gca gcg gcg tca aaa gtg aag 16455  
 Tyr Gln Asp Val Leu Lys Glu Val Lys Ala Ala Ser Lys Val Lys  
 1245 1250 1255

gct aac ttg cta tcc gta gag gaa gct tgc agc ctg acg ccc cca cac 16503  
 Ala Asn Leu Leu Ser Val Glu Glu Ala Cys Ser Leu Thr Pro Pro His  
 1260 1265 1270 1275

tca gcc aaa tcc aag ttt ggt tat ggg gca aaa gac gtc cgt tgc cat 16551  
 Ser Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Cys His  
 1280 1285 1290

gcc aga aag gcc gta acc cac atc aac tcc gtg tgg aaa gac ctt ctg 16599  
 Ala Arg Lys Ala Val Thr His Ile Asn Ser Val Trp Lys Asp Leu Leu  
 1295 1300 1305

gaa gac aat gta aca cca ata gac act acc atc atg gct aag aac gag 16647  
 Glu Asp Asn Val Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu  
 1310 1315 1320

gtt ttc tgc gtt cag cct gag aag ggg ggt cgt aag cca gct cgt ctc 16695  
 Val Phe Cys Val Gln Pro Glu Lys Gly Arg Lys Pro Ala Arg Leu  
 1325 1330 1335

atc gtg ttc ccc gat ctg ggc gtc cgc gtg tgc gaa aag atg gct ttg 16743  
 Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu  
 1340 1345 1350 1355

tac gac gtg gtt aca aag ctc acc ttg gcc gtg atg gga agc tcc tac 16791  
 Tyr Asp Val Val Thr Lys Leu Pro Leu Ala Val Met Gly Ser Ser Tyr  
 1360 1365 1370

gga ttc caa tac tca cca gga cag cgg gtt gaa ttc ctc gtg caa gcg 16839

DRAFT

Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Gln Ala			
1375	1380	1385	
tgg aag tcc aag aaa acc cca atg ggg ttc tcg tat gat acc cgc tgc			16887
Trp Lys Ser Lys Lys Thr Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys			
1390	1395	1400	
ttt gac tcc aca gtc act gag agc gac atc cgt acg gag gag gca atc			16935
Phe Asp Ser Thr Val Thr Glu Ser Asp Ile Arg Thr Glu Glu Ala Ile			
1405	1410	1415	
tac caa tgt tgt gac ctc gac ccc caa gcc cgc gtg gcc atc aag tcc			16983
Tyr Gln Cys Cys Asp Leu Asp Pro Gln Ala Arg Val Ala Ile Lys Ser			
1420	1425	1430	1435
ctc acc gag agg ctt tat gtt ggg ggc cct ctt acc aat tca agg ggg			17031
Leu Thr Glu Arg Leu Tyr Val Gly Gly Pro Leu Thr Asn Ser Arg Gly			
1440	1445	1450	
gag aac tgc ggc tat cgc agg tgc cgc gcg agc ggc gta ctg aca act			17079
Glu Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr			
1455	1460	1465	
agc tgt ggt aac acc ctc act tgc tac atc aag gcc cggt gca gcc tgt			17127
Ser Cys Gly Asn Thr Leu Thr Cys Tyr Ile Lys Ala Arg Ala Ala Cys			
1470	1475	1480	
cga gcc gca ggg ctc cag gac tgc acc atg ctc gtgt tgt ggc gac gac			17175
Arg Ala Ala Gly Leu Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp			
1485	1490	1495	
tta gtc gtt atc tgt gaa agc gcg ggg gtc cag gag gac gcg gcg agc			17223
Leu Val Val Ile Cys Glu Ser Ala Gly Val Gln Glu Asp Ala Ala Ser			
1500	1505	1510	1515
ctg aga gcc ttc acg gag gct atg acc agg tac tcc gcc ccc cct ggg			17271
Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly			
1520	1525	1530	
gac ccc cca caa cca gaa tac gac ttg gag ctc ata aca tca tgc tcc			17319
Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser			
1535	1540	1545	
tcc aac gtg tca gtc gcc cac gac ggc gct gga aag agg gtc tac tac			17367
Ser Asn Val Ser Val Ala His Asp Gly Ala Gly Lys Arg Val Tyr Tyr			
1550	1555	1560	
ctc acc cgt gac cct aca acc ccc ctc gcg aga gct gcg tgg gag aca			17415
Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr			
1565	1570	1575	
gca aga cac act cca gtc aat tcc tgg cta ggc aac ata atc atg ttt			17463
Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Phe			
1580	1585	1590	1595
gcc ccc aca ctg tgg gcg agg atg ata ctg atg acc cat ttc ttt agc			17511
Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Ser			

1600                    1605                    1610

gtc ctt ata gcc agg gac cag ctt gaa cag gcc ctc gat tgc gag atc      17559  
 Val Leu Ile Ala Arg Asp Gln Leu Glu Gln Ala Leu Asp Cys Glu Ile  
 1615                    1620                    1625

tac ggg gcc tgc tac tcc ata gaa cca ctg gat cta cct cca atc att      17607  
 Tyr Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro Pro Ile Ile  
 1630                    1635                    1640

caa aga ctc cat ggc ctc agc gca ttt tca ctc cac agt tac tct cca      17655  
 Gln Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro  
 1645                    1650                    1655

ggt gaa atc aat agg gtg gcc gca tgc ctc aga aaa ctt ggg gta ccg      17703  
 Gly Glu Ile Asn Arg Val Ala Ala Cys Leu Arg Lys Leu Gly Val Pro  
 1660                    1665                    1670                    1675

ccc ttg cga gct tgg aga cac cgg gcc cgg agc gtc cgc gct agg ctt      17751  
 Pro Leu Arg Ala Trp Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu  
 1680                    1685                    1690

ctg gcc aga gga ggc agg gct gcc ata tgt ggc aag tac ctc ttc aac      17799  
 Leu Ala Arg Gly Gly Arg Ala Ala Ile Cys Gly Lys Tyr Leu Phe Asn  
 1695                    1700                    1705

tgg gca gta aga aca aag ctc aaa ctc act cca ata gcg gcc gct ggc      17847  
 Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Ala Ala Gly  
 1710                    1715                    1720

cag ctg gac ttg tcc ggc tgg ttc acg gct ggc tac agc ggg gga gac      17895  
 Gln Leu Asp Leu Ser Gly Trp Phe Thr Ala Gly Tyr Ser Gly Gly Asp  
 1725                    1730                    1735

att tat cac agc gtg tct cat gcc cgg ccc cgc tgg atc tgg ttt tgc      17943  
 Ile Tyr His Ser Val Ser His Ala Arg Pro Arg Trp Ile Trp Phe Cys  
 1740                    1745                    1750                    1755

cta ctc ctg ctt gct gca ggg gta ggc atc tac ctc ctc ccc aac cga      17991  
 Leu Leu Leu Ala Ala Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg  
 1760                    1765                    1770

atg agc acg aat cct aaa cct caa aga aag acc aaa cgt aac acc aac      18039  
 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn  
 1775                    1780                    1785

cgg cgg ccg cag gac gtc aag ttc cgg ggt ggc ggt cag atc gtt ggt      18087  
 Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gln Ile Val Gly  
 1790                    1795                    1800

gga gtt tac ttg ttg ccg cgc agg ggc cct aga ttg ggt gtg cgc gcg      18135  
 Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala  
 1805                    1810                    1815

acg aga aag act tcc gag cgg tcg caa cct cga ggt aga cgt cag cct      18183  
 Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro  
 1820                    1825                    1830                    1835

atc ccc aag gct cgt cgcc gag ggc agg acc tgg gct cag ccc ggg 18231  
 Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly  
     1840                1845                1850

tac cct tgg ccc ctc tat ggc aat gag ggc tgc ggg tgg gcg gga tgg 18279  
 Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp  
     1855                1860                1865

ctc ctg tct ccc cgt tct cgg cct agc tgg ggc ccc aca gac ccc 18327  
 Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro  
     1870                1875                1880

cgcg cgt agg tcg cgc aat ttg ggt aag gtc atc gat acc ctt acg tgc 18375  
 Arg Arg Arg Ser Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys  
     1885                1890                1895

ggc ttc gcc gac ctc atg ggg tac ata ccg ctc gtc ggc ggc cct ctt 18423  
 Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu  
     1900                1905                1910                1915

gga ggc gct gcc agg gcc taatagtcga ctttggccc actgtacttt 18471  
 Gly Gly Ala Ala Arg Ala  
     1920

tagctcgtaaaaatacatatactttca tttctccgtaaacaacatgt tttccatgt 18531  
 aatatccttttctattttc gttccgttac caactttaca catactttat atagctattc 18591  
 acttctatac actaaaaaac taagacaatt ttaatttgc tgcctgccat atttcaattt 18651  
 gttataaatt cctataattt atcctattag tagctaaaaa aagatgaatg tgaatcgaat 18711  
 cctaagagaa ttggatctga tccacaggac gggtgtggc gccatgatcg cgtagtcgat 18771  
 agtggctcca agtagcgaag cgagcaggac tggcgccgg ccaaagcggt cggacagtgc 18831  
 tccgagaacg ggtgcgcata gaaattgcat caacgcata agcgctagca gcacgccata 18891  
 gtgactggcg atgctgtcg aatggacgat atccgcaag aggccggca gtaccggcat 18951  
 aaccaagcct atgcctacag catccagggt gacggtgccg aggatgacga tgagcgcatt 19011  
 gttagatttc atacacggtg cctgactgctttagcaattt aactgtgata aactaccgca 19071  
 ttaaagcttttctttccaa tttttttttt ttcgtcatta taaaaatcat tacgaccgag 19131  
 attccccgggt aataactgat ataattaaat tgaagctcta atttgtgagt ttagtataca 19191  
 tgcatttact tataatacag ttttttagtt ttgctggccg catcttctca aatatgcttc 19251  
 ccagcctgct tttctgttaac gttcacccctc taccttagca tcccttcctt tgcaaata 19311  
 tcctcttcca acaataataa tgttagatcc tgttagagacc acatcatcca cgggtctata 19371  
 ctgttgaccc aatgcgtctc cttgtcatac taaacccaca cgggtgtca taatcaacca 19431  
 atcgtaacct tcatcttcc acccatgtc tcttgagca ataaagccga taacaaaatc 19491

tttgcgttc ttgcgtatgt caacagtacc ctttgttat tctccagtag ataggagcc 19551  
cttgcgtgac aattctgcta acatcaaaag gcctctaggt tcctttgtta cttcttctgc 19611  
cgccctgcttc aaaccgctaa caatacctgg gcccaccaca ccgtgtgcat tcgtaatgtc 19671  
tgcccattct gctattctgt atacacccgc agagtactgc aatttgcgt tattaccaat 19731  
gtcagcaaat tttctgtctt cgaagagtaa aaaattgtac ttggcggata atgcctttag 19791  
cggttaact gtgcctcca tggaaaaatc agtcaagata tccacatgtg ttttagtaa 19851  
acaaatttg ggacctaattt cttcaactaa ctccagtaat tccttgggg tacgaacatc 19911  
caatgaagca cacaagttt tttgttttc gtgcgtgata taaaatagct tggcagcaac 19971  
aggacttagga tgagtagcag cacgttcctt atatgtagct ttcgacatga tttatctcg 20031  
tttcctgcag gttttgttc tgtgcgttg ggttaagaat actgggcaat ttcatgttcc 20091  
ttcaacacta catatgcgta tatataccaa tctaagtctg tgcccttcc ttcgttcttc 20151  
cttctgttcg gagattaccg aatcaaaaaa attcaagga aaccgaaatc aaaaaaaaga 20211  
ataaaaaaaaaa aatgatgaat tgaaaagctt atcgat 20247

<210> 19  
<211> 1921  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
pd.delta.NS3NS5.pj.core150

<400> 19  
Met Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro  
1 5 10 15  
Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His  
20 25 30  
Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly  
35 40 45  
Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly  
50 55 60  
Cys Ser Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser  
65 70 75 80  
Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Ala  
85 90 95  
Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala Thr Ala Thr Pro Pro  
100 105 110

Gly Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser  
115 120 125

Thr Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val  
130 135 140

Ile Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Lys Lys Cys  
145 150 155 160

Asp Glu Leu Ala Ala Lys Leu Val Ala Leu Gly Ile Asn Ala Val Ala  
165 170 175

Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val  
180 185 190

Val Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe  
195 200 205

Asp Ser Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe  
210 215 220

Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr Ile Thr Leu Pro Gln Asp  
225 230 235 240

Ala Val Ser Arg Thr Gln Arg Arg Gly Arg Thr Gly Arg Gly Lys Pro  
245 250 255

Gly Ile Tyr Arg Phe Val Ala Pro Gly Glu Arg Pro Ser Gly Met Phe  
260 265 270

Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp Ala Gly Cys Ala Trp Tyr  
275 280 285

Glu Leu Thr Pro Ala Glu Thr Thr Val Arg Leu Arg Ala Tyr Met Asn  
290 295 300

Thr Pro Gly Leu Pro Val Cys Gln Asp His Leu Glu Phe Trp Glu Gly  
305 310 315 320

Val Phe Thr Gly Leu Thr His Ile Asp Ala His Phe Leu Ser Gln Thr  
325 330 335

Lys Gln Ser Gly Glu Asn Leu Pro Tyr Leu Val Ala Tyr Gln Ala Thr  
340 345 350

Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser Trp Asp Gln Met Trp  
355 360 365

Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His Gly Pro Thr Pro Leu  
370 375 380

Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Ile Thr Leu Thr His Pro  
385 390 395 400

Val Thr Lys Tyr Ile Met Thr Cys Met Ser Ala Asp Leu Glu Val Val  
405 410 415

Thr Ser Thr Trp Val Leu Val Gly Gly Val Leu Ala Ala Leu Ala Ala  
 420 425 430

Tyr Cys Leu Ser Thr Gly Cys Val Val Ile Val Gly Arg Val Val Leu  
 435 440 445

Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu Val Leu Tyr Arg Glu  
 450 455 460

Phe Asp Glu Met Glu Glu Cys Ser Gln His Leu Pro Tyr Ile Glu Gln  
 465 470 475 480

Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys Ala Leu Gly Leu Leu  
 485 490 495

Gln Thr Ala Ser Arg Gln Ala Glu Val Ile Ala Pro Ala Val Gln Thr  
 500 505 510

Asn Trp Gln Lys Leu Glu Thr Phe Trp Ala Lys His Met Trp Asn Phe  
 515 520 525

Ile Ser Gly Ile Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn  
 530 535 540

Pro Ala Ile Ala Ser Leu Met Ala Phe Thr Ala Ala Val Thr Ser Pro  
 545 550 555 560

Leu Thr Thr Ser Gln Thr Leu Leu Phe Asn Ile Leu Gly Gly Trp Val  
 565 570 575

Ala Ala Gln Leu Ala Ala Pro Gly Ala Ala Thr Ala Phe Val Gly Ala  
 580 585 590

Gly Leu Ala Gly Ala Ala Ile Gly Ser Val Gly Leu Gly Lys Val Leu  
 595 600 605

Ile Asp Ile Leu Ala Gly Tyr Gly Ala Gly Val Ala Gly Ala Leu Val  
 610 615 620

Ala Phe Lys Ile Met Ser Gly Glu Val Pro Ser Thr Glu Asp Leu Val  
 625 630 635 640

Asn Leu Leu Pro Ala Ile Leu Ser Pro Gly Ala Leu Val Val Gly Val  
 645 650 655

Val Cys Ala Ala Ile Leu Arg Arg His Val Gly Pro Gly Glu Gly Ala  
 660 665 670

Val Gln Trp Met Asn Arg Leu Ile Ala Phe Ala Ser Arg Gly Asn His  
 675 680 685

Val Ser Pro Thr His Tyr Val Pro Glu Ser Asp Ala Ala Ala Arg Val  
 690 695 700

Thr Ala Ile Leu Ser Ser Leu Thr Val Thr Gln Leu Leu Arg Arg Leu  
 705 710 715 720

His Gln Trp Ile Ser Ser Glu Cys Thr Thr Pro Cys Ser Gly Ser Trp  
725 730 735

Leu Arg Asp Ile Trp Asp Trp Ile Cys Glu Val Leu Ser Asp Phe Lys  
740 745 750

Thr Trp Leu Lys Ala Lys Leu Met Pro Gln Leu Pro Gly Ile Pro Phe  
755 760 765

Val Ser Cys Gln Arg Gly Tyr Lys Gly Val Trp Arg Gly Asp Gly Ile  
770 775 780

Met His Thr Arg Cys His Cys Gly Ala Glu Ile Thr Gly His Val Lys  
785 790 795 800

Asn Gly Thr Met Arg Ile Val Gly Pro Arg Thr Cys Arg Asn Met Trp  
805 810 815

Ser Gly Thr Phe Pro Ile Asn Ala Tyr Thr Thr Gly Pro Cys Thr Pro  
820 825 830

Leu Pro Ala Pro Asn Tyr Thr Phe Ala Leu Trp Arg Val Ser Ala Glu  
835 840 845

Glu Tyr Val Glu Ile Arg Gln Val Gly Asp Phe His Tyr Val Thr Gly  
850 855 860

Met Thr Thr Asp Asn Leu Lys Cys Pro Cys Gln Val Pro Ser Pro Glu  
865 870 875 880

Phe Phe Thr Glu Leu Asp Gly Val Arg Leu His Arg Phe Ala Pro Pro  
885 890 895

Cys Lys Pro Leu Leu Arg Glu Glu Val Ser Phe Arg Val Gly Leu His  
900 905 910

Glu Tyr Pro Val Gly Ser Gln Leu Pro Cys Glu Pro Glu Pro Asp Val  
915 920 925

Ala Val Leu Thr Ser Met Leu Thr Asp Pro Ser His Ile Thr Ala Glu  
930 935 940

Ala Ala Gly Arg Arg Leu Ala Arg Gly Ser Pro Pro Ser Val Ala Ser  
945 950 955 960

Ser Ser Ala Ser Gln Leu Ser Ala Pro Ser Leu Lys Ala Thr Cys Thr  
965 970 975

Ala Asn His Asp Ser Pro Asp Ala Glu Leu Ile Glu Ala Asn Leu Leu  
980 985 990

Trp Arg Gln Glu Met Gly Gly Asn Ile Thr Arg Val Glu Ser Glu Asn  
995 1000 1005

Lys Val Val Ile Leu Asp Ser Phe Asp Pro Leu Val Ala Glu Glu Asp  
1010 1015 1020

Glu Arg Glu Ile Ser Val Pro Ala Glu Ile Leu Arg Lys Ser Arg Arg  
025 1030 1035 1040

Phe Ala Gln Ala Leu Pro Val Trp Ala Arg Pro Asp Tyr Asn Pro Pro  
1045 1050 1055

Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr Glu Pro Pro Val Val His  
1060 1065 1070

Gly Cys Pro Leu Pro Pro Lys Ser Pro Pro Val Pro Pro Pro Arg  
1075 1080 1085

Lys Lys Arg Thr Val Val Leu Thr Glu Ser Thr Leu Ser Thr Ala Leu  
1090 1095 1100

Ala Glu Leu Ala Thr Arg Ser Phe Gly Ser Ser Ser Thr Ser Gly Ile  
1095 1110 1115 1120

Thr Gly Asp Asn Thr Thr Ser Ser Glu Pro Ala Pro Ser Gly Cys  
1125 1130 1135

Pro Pro Asp Ser Asp Ala Glu Ser Tyr Ser Ser Met Pro Pro Leu Glu  
1140 1145 1150

Gly Glu Pro Gly Asp Pro Asp Leu Ser Asp Gly Ser Trp Ser Thr Val  
1155 1160 1165

Ser Ser Glu Ala Asn Ala Glu Asp Val Val Cys Cys Ser Met Ser Tyr  
1170 1175 1180

Ser Trp Thr Gly Ala Leu Val Thr Pro Cys Ala Ala Glu Glu Gln Lys  
1185 1190 1195 1200

Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu Arg His His Asn Leu  
1205 1210 1215

Val Tyr Ser Thr Thr Ser Arg Ser Ala Cys Gln Arg Gln Lys Lys Val  
1220 1225 1230

Thr Phe Asp Arg Leu Gln Val Leu Asp Ser His Tyr Gln Asp Val Leu  
1235 1240 1245

Lys Glu Val Lys Ala Ala Ala Ser Lys Val Lys Ala Asn Leu Leu Ser  
1250 1255 1260

Val Glu Glu Ala Cys Ser Leu Thr Pro Pro His Ser Ala Lys Ser Lys  
1265 1270 1275 1280

Phe Gly Tyr Gly Ala Lys Asp Val Arg Cys His Ala Arg Lys Ala Val  
1285 1290 1295

Thr His Ile Asn Ser Val Trp Lys Asp Leu Leu Glu Asp Asn Val Thr  
1300 1305 1310

Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu Val Phe Cys Val Gln  
1315 1320 1325

Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu Ile Val Phe Pro Asp  
 1330 1335 1340

Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu Tyr Asp Val Val Thr  
 345 1350 1355 1360

Lys Leu Pro Leu Ala Val Met Gly Ser Ser Tyr Gly Phe Gln Tyr Ser  
 1365 1370 1375

Pro Gly Gln Arg Val Glu Phe Leu Val Gln Ala Trp Lys Ser Lys Lys  
 1380 1385 1390

Thr Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe Asp Ser Thr Val  
 1395 1400 1405

Thr Glu Ser Asp Ile Arg Thr Glu Glu Ala Ile Tyr Gln Cys Cys Asp  
 1410 1415 1420

Leu Asp Pro Gln Ala Arg Val Ala Ile Lys Ser Leu Thr Glu Arg Leu  
 425 1430 1435 1440

Tyr Val Gly Gly Pro Leu Thr Asn Ser Arg Gly Glu Asn Cys Gly Tyr  
 1445 1450 1455

Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr Ser Cys Gly Asn Thr  
 1460 1465 1470

Leu Thr Cys Tyr Ile Lys Ala Arg Ala Cys Arg Ala Ala Gly Leu  
 1475 1480 1485

Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp Leu Val Val Ile Cys  
 1490 1495 1500

Glu Ser Ala Gly Val Gln Glu Asp Ala Ala Ser Leu Arg Ala Phe Thr  
 505 1510 1515 1520

Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp Pro Pro Gln Pro  
 1525 1530 1535

Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser Asn Val Ser Val  
 1540 1545 1550

Ala His Asp Gly Ala Gly Lys Arg Val Tyr Tyr Leu Thr Arg Asp Pro  
 1555 1560 1565

Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr Ala Arg His Thr Pro  
 1570 1575 1580

Val Asn Ser Trp Leu Gly Asn Ile Ile Met Phe Ala Pro Thr Leu Trp  
 585 1590 1595 1600

Ala Arg Met Ile Leu Met Thr His Phe Phe Ser Val Leu Ile Ala Arg  
 1605 1610 1615

Asp Gln Leu Glu Gln Ala Leu Asp Cys Glu Ile Tyr Gly Ala Cys Tyr  
 1620 1625 1630

Ser Ile Glu Pro Leu Asp Leu Pro Pro Ile Ile Gln Arg Leu His Gly  
1635 1640 1645

Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro Gly Glu Ile Asn Arg  
1650 1655 1660

Val Ala Ala Cys Leu Arg Lys Leu Gly Val Pro Pro Leu Arg Ala Trp  
665 1670 1675 1680

Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu Leu Ala Arg Gly Gly  
1685 1690 1695

Arg Ala Ala Ile Cys Gly Lys Tyr Leu Phe Asn Trp Ala Val Arg Thr  
1700 1705 1710

Lys Leu Lys Leu Thr Pro Ile Ala Ala Gly Gln Leu Asp Leu Ser  
1715 1720 1725

Gly Trp Phe Thr Ala Gly Tyr Ser Gly Gly Asp Ile Tyr His Ser Val  
1730 1735 1740

Ser His Ala Arg Pro Arg Trp Ile Trp Phe Cys Leu Leu Leu Ala  
745 1750 1755 1760

Ala Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg Met Ser Thr Asn Pro  
1765 1770 1775

Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Gln Asp  
1780 1785 1790

Val Lys Phe Pro Gly Gly Gln Ile Val Gly Val Tyr Leu Leu  
1795 1800 1805

Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala Thr Arg Lys Thr Ser  
1810 1815 1820

Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro Ile Pro Lys Ala Arg  
825 1830 1835 1840

Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly Tyr Pro Trp Pro Leu  
1845 1850 1855

Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp Leu Leu Ser Pro Arg  
1860 1865 1870

Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro Arg Arg Ser Arg  
1875 1880 1885

Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu  
1890 1895 1900

Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg  
905 1910 1915 1920

Ala